12-14-2011

Investigating block mask lithography variation using finite-difference time-domain simulation

Daiji Kawamura

INVESTIGATING BLOCK MASK LITHOGRAPHY VARIATION USING FINITE-DIFFERENCE TIME-DOMAIN SIMULATION

By

Daiji Kawamura

A thesis submitted in
Partial fulfillment of the requirements for the degree of
Master of Science in
Microelectronic Engineering

Approved by:

Dr. Karl Hirschman _____ Date: _____
(Thesis Advisor)

Dr. Dale Ewbank _____ Date: _____
(Committee Member)

Dr. Robert Pearson _____ Date: _____
(Program Director)

Dr. Sohail Dianat _____ Date: _____
(EME Department Head)

Electrical and Microelectronic Engineering Department
Kate Gleason College of Engineering
Rochester Institute of Technology
December 14th, 2011

INVESTIGATING BLOCK MASK LITHOGRAPHY VARIATION

USING FINITE-DIFFERENCE TIME-DOMAIN SIMULATION

By

Daiji Kawamura

I, Daiji Kawamura, hereby grant permission to the Wallace Memorial Library of the Rochester Institute of Technology to reproduce this document in whole or in part that any reproduction will not be for commercial use or profit.

_____ _____

Daiji Kawamura December 14th, 2012

ABSTRACT

Simulation work has long been realized as a method for analyzing semiconductor processing expediently and cost-effectively. As technology advancements strive to meet increasingly stringent parameter constraints, difficult issues arise. In this paper, challenges in block mask lithography will be discussed with the aid of using simulation packages developed by Panoramic Technology®. Halo formation utilizes a 20-30° tilt-angle implantation [1]. The block mask defines the geometries of the resist opening to allow implantation of atoms to extend into the channel region. Due to designed resolution scaling and tolerance in conjunction with substrate topography, there can be undesired influence on the electrical device characteristics due to block variations. Although the block mask pattern definition is relatively simple, additional investigation is required to understand the sensitivities that drive the implant resist CD variation.

In this study, block mask measurements processed using 248 nm and 193 nm illumination sources were used to calibrate the simulation work. Addition of optical proximity correction (OPC) and wafer topography geometry parameters have been shown to improve modeling capabilities. The modeling work was also able to show the benefits of a developable bottom anti-reflection coating (dBARC) process over a single layer resist (SLR) process in the resist intensity profiles as gate pitch is decreased. The goal of this work was to develop an accurate simulation model that characterizes the lithographic performance needed to support the transition into future technology nodes.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS AND SYMBOLS

α	absorption coefficient
ASIC	application specific integrated circuit
BARC	bottom anti-reflective coating
BCP	block copolymer
c	speed of light
CD	critical dimension
COO	cost of ownership
DBARC	developable bottom anti-reflective coating
DIBL	drain-induced barrier lowering
DNQ	diazonaphthoquinine
DRAM	dynamic random access memory
EM	electromagnetic
emf	electromotive force
EUV	extreme ultra-violet
EUVL	EUV lithography
FDTD	finite-difference time-domain
h	Planck's constant
κ	index of refraction – imaginary component
ITRS	International Technology Roadmap for Semiconductors
LER	line edge roughness
ML	maskless lithography
MPU	micro processing unit
NA	numerical aperture
NIL	nanoimprint lithography
OPC	optical proximity correction
PAC	photoactive compound
PAG	photoacid generator
PC	polysilicon conductors
PEB	post exposure bake
RET	resolution enhancement technique
RMSE	root-mean-square error
σ	partial coherence
S/D	source-drain
SCE	short channel effect
SEM	scanning electron micrograph
SLR	single layer resist
TBS	t-BOC polystyrene
TEMPEST	time-domain electromagnetic massively parallel evaluation of scattering from topography
λ	wavelength

ACKNOWLEDGEMENTS

I would like to thank God for the innumerable blessings in my life and the amazing people who have helped and guided me to complete my thesis document. To Matt Colburn, thank you for your continual guidance and helping me establish a publishable and exciting project. To Dr. Karl Hirschman, thank you for unwavering support and helping me grow in my educational career over the years. To Dr. Dale Ewbank and Dr. Robert Pearson, thank you for your feedback and help to fine-tune my document, even through the holidays. To Steve Holmes, Greg McIntyre, Narasim Kanike, Wally Carpenter, and my IBM colleagues, thank you for providing insight and continuingly helping me through this project. To the RIT student, staff and support, thank you for your cooperative work and equipping me with the necessary knowledge to pursuing my degree. To my family, thank you for always being there and supporting me all my life. To my friends, thank you for believing in me and helping me accomplish this work. I could not have achieved this work without all of your support, so thank you sincerely for helping me achieve my dream and starting a great future.

CHAPTER 1

MOTIVATION

Understanding the physical nature of light at the device level is crucial to lithography processing. As devices begin to scale, topographical features can limit the amount of exposure in certain regions. The project presented in this paper will investigate two main areas of lithography that are of current interest in CMOS fabrication. The first issue is variation of implant block mask critical dimension (CD) at a fixed gate pitch. In block mask lithography processing, there are issues concerning resist imaging that lead to CD variation. Effects of light scattering from underlying polysilicon-conductor (PC) topographic features can cause corner rounding of block mask edges, which has been linked to variation in the transistor threshold voltage. An example illustration of this effect is shown in Figure 1.

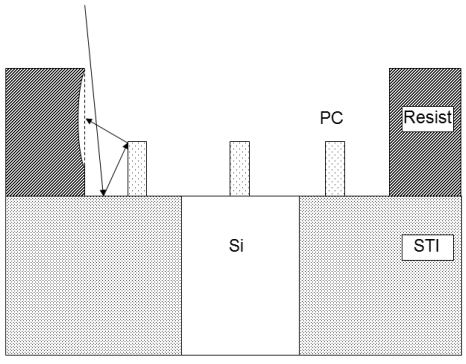


Figure 1. An illustration showing wafer topographical effects on resist profile. An incoming electromagnetic ray may reflect off a corner of a polysilicon conductor (PC) and cause resist exposure on the edge. The term PC topography is introduced to represent the entire wafer topography due to gate-like structures. Refer to Chapter 4 for further explanation.

The connection between block mask CD variation and the device electrical characteristics is captured in Figure 2, which shows the threshold voltage (V_T) response over different CD targets. The observed non-random shift in V_T at specific test macro ID sites is caused by an influence of the block mask resist feature on the angled implant into the active area. Thus it is beneficial to understand the mechanism of CD variation and model CD deviation from target values in the block mask resist image.

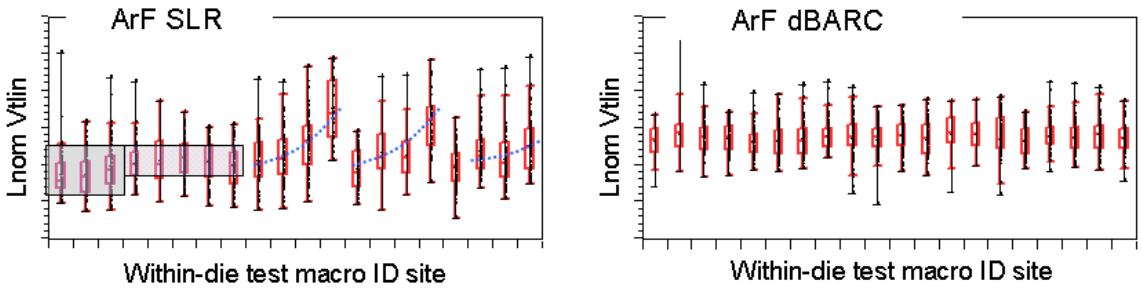


Figure 2. Variability chart for L_{nom} V_{Tlin} of a SLR processing (left) threshold voltage with various macro dimensions are shown to have an upward trend in the latter macro dimensions. The use of a dBARC with thin resist (right) seems to show an improvement in device behavior [2].

The second issue is variation in the resist intensity profiles due to a dependence on the underlying PC pitch. As the technology node implement decreasing gate pitches, the propagation of light becomes affected such that the light intensity becomes minimal at the edges of the PC. In standard processing at the current technology node with positive single-layer resist, this can result in residual resist at the feature corners. Understanding this interaction of PC pitch variation and light intensity profiles is critical. Three-dimensional simulation capabilities for these issues are extremely beneficial to predict and analyze the imaging outcome.

CHAPTER 2

INTRODUCTION

Semiconductor technology is constantly in pursuit of economic and performance scaling improvement. In the semiconductor industry, Moore's law has predicted the remarkable growth in the industry for over 40 years. Moore's law states that the number of transistors on a chip doubles every two years [3]. The law itself is a simple statement but the actual requirements that are needed to fulfill this trend are complex.

2.1 TECHNOLOGY NODE AND CHALLENGES

In an industry that's driven by Moore's law to maintain the pace of technological advancements, lithography technology has been a determining factor in sustaining the device miniaturization trend. Historically, in the mid-1980's, many believed that optical lithography was limited to resolving dimensions larger than one micron [4]. However, this limitation has long been overcome enabling conventional lithography techniques to imprint technology nodes smaller than 32 nm [5]. Semiconductor technology nodes are defined by the critical dimension (CD) given by the Rayleigh CD criterion as shown in the equation:

$$CD = k_1 \frac{\lambda}{NA}. \quad (1)$$

The above equation shows that CD resolution can be enhanced by shorter wavelengths (λ), higher numerical apertures (NA), and reducing k_1 values. Mathematically, this is a

simple relationship. In reality, implementing finer resolution defined by the Rayleigh CD criterion gives rise to several challenges. Extreme ultraviolet (EUV) lithography, or EUVL, is currently the leading next-generation lithography because of its low wavelength exposure at 13.5 nm. However, difficult challenges such as efficient and high power light sources, high precision reflective optics, and low-defect masks, continue to keep EUVL in the developmental stages [6]. Fortunately, methods such as immersion lithography and double-patterning have allowed for continued device scaling utilizing the 193 nm wavelength.

2.2 LITHOGRAPHY ITRS ROADMAP

The International Technology Roadmap for Semiconductors (ITRS) gives a thorough trend analysis of future technology requirements [7]. A graphical trend of flash and DRAM ½ pitch requirements are shown in Figure 3. Although the cited values shown in Table 1 are based on current best estimates, the ITRS gives a helpful guideline

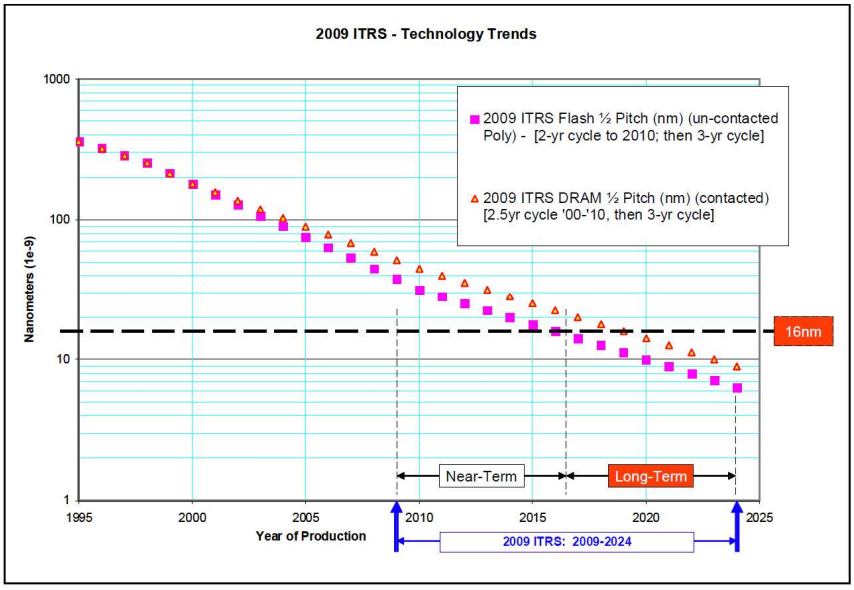


Figure 3. An overall roadmap technology trend of DRAM and Flash memory half pitch based on the ITRS estimates [8].

for the semiconductor industry technological advancements. The simulation parameters in Chapter 4 were based on the MPU/ASIC ½ pitch values.

Table 1. Targeted values of technology requirements of key lithography-related products based on the ITRS [8].

Near-term Years								
YEAR OF PRODUCTION:	*2009*	*2010*	*2011*	*2012*	*2013*	*2014*	*2015*	*2016*
Flash Uncontacted Poly Si ½ Pitch (nm)	38	32	28	25	23	20	18	15.9
DRAM stagger-contacted Metal 1 (M1) ½ Pitch (nm)	52	45	40	36	32	28	25	22.5
MPU/ASIC stagger-contacted Metal 1 (M1) ½ Pitch (nm)	54	45	38	32	27	24	21	18.9
MPU Printed Gate Length (nm)	47	41	35	31	28	25	22	19.8

Long-term Years								
YEAR OF PRODUCTION:	*2017*	*2018*	*2019*	*2020*	*2021*	*2022*	*2023*	*2024*
Flash Uncontacted Poly Si ½ Pitch (nm)	14.2	12.6	11.3	10	8.9	8.0	7.1	6.3
DRAM stagger-contacted Metal 1 (M1) ½ Pitch (nm)	20	17.9	15.9	14.2	12.6	11.3	10	8.9
MPU/ASIC stagger-contacted Metal 1 (M1) ½ Pitch (nm)	16.9	15	13.4	11.9	10.6	9.5	8.4	7.5
MPU Printed Gate Length (nm)	17.7	15.7	14	12.5	11.1	9.9	8.8	7.9

In the current development of lithographic techniques, engineers have the challenge of considering two tasks simultaneously so that economical benefits are optimized [9]. The two tasks are:

1. Develop and improve methods to extend the incumbent optical lithography technology.
2. Develop and investigate alternative methods for the next generation lithography technologies (NGL).

In the first task there are numerous research studies from the most complex, including the development of immersion resist components [10], solid immersion lenses [11], and multiple exposure techniques [12], to the most straightforward such as utilizing developable bottom anti-reflective coatings (dBARC) [13]. The dBARC processing

method is a primary focus of this paper because of its benefit seen in block mask lithography. In the second task, there are technologies such as x-ray [14], electron beam/maskless lithography [15]-[16], and EUVL [17]. EUVL is currently the most promising candidate for NGL to meet pattern formation requirements with resolution better than 20 nm [9]. The combination of optimizing innovative methods and developing new technologies makes the lithography industry both challenging and exciting.

2.3 SHORT CHANNEL EFFECTS AND ION IMPLANTATION

The importance of ion implantation and synthesis of materials [18] must be acknowledged. It is the processing of ion implantation that allows introducing dopant atoms into silicon regions. It is the combination of lithography and ion implantation that allows for selectively altering the conductivity of the semiconductor. As device dimensions shrink, there are undesirable short-channel effects that can be addressed by specific implants.

Channel length reduction can give rise to problematic issues of short channel effects (SCEs). As depicted in Figure 4(a), Muller and Kamins show how SCEs are depicted in a cross sectional view of both long (dotted lines) and short channel (solid lines) devices [19]. Some common SCEs in advanced CMOS fabrication are punch-through, drain-induced barrier lowering, and hot carrier effects. When the gate channel reduction is significant, the underlying depletion regions of the source and drain (S/D) extend into the channel region. The drain bias can cause a lowering of the barrier potential at the source end, as shown in Figure 4(b), which results in poor gate control.

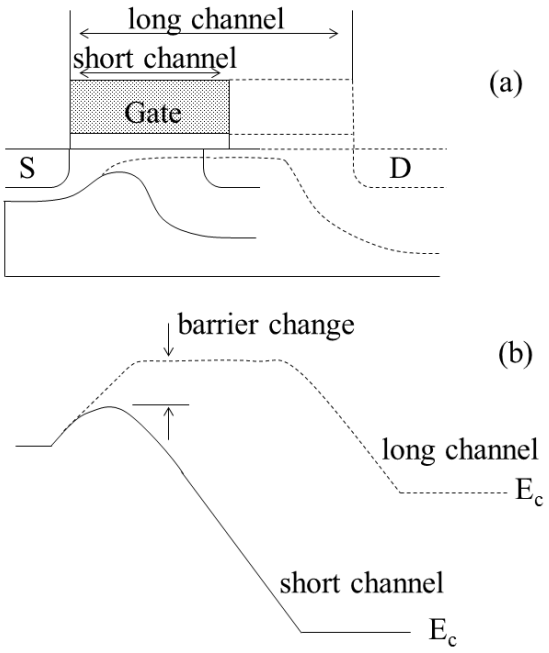


Figure 4. An illustration of short channel effect in a (a) cross sectional view of the gate, source, and drain and (b) conduction band energy along the surface. The long-channel condition is represented by the dotted lines while the solid line represents the short-channel condition [19].

This is known as the drain-induced barrier lowering (DIBL) because the decreased barrier height allows easier carrier diffusion that negatively impacts gate control of the current. A punch-through effect can occur at extreme conditions where the S/D depletion region extends into the channel region, preventing the gate control of the current [18]. Ion implantation has been successfully used to counter these short-channel effects.

Punch-through-stop Implant

A punch-through-stop implant is commonly used to reduce parasitic current paths between the S/D. By increasing the bulk doping level of the silicon, it prevents the S/D depletion region intrusion into the gate region. The abrupt doping profile below the S/D depth confines the current path and avoids current leakage. Arsenic or antimony

dopants are typically used for PMOS transistors while indium is used for NMOS transistors. These dopant species have a slow diffusing characteristic that allow an abrupt doping profile [18].

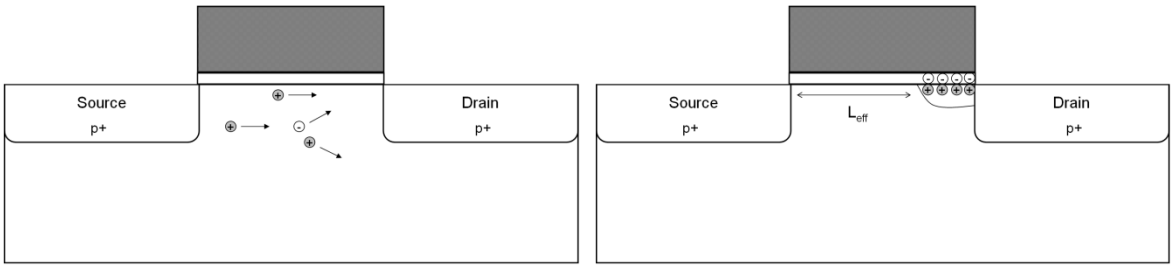


Figure 5. An illustration showing hot electron effects that can cause a shortening of gate channel length to L_{eff} [18].

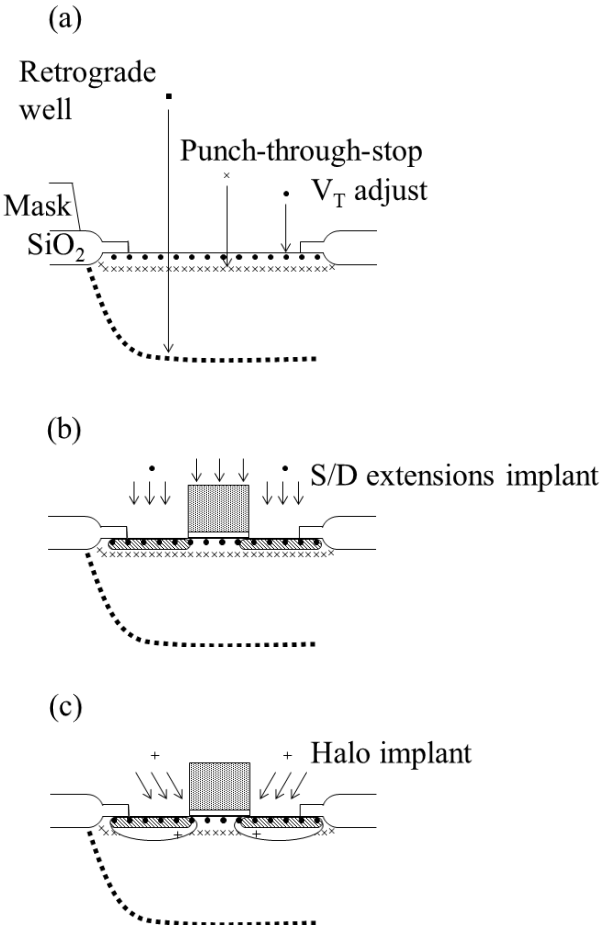
Extension (LDD) Formation

S/D extensions, or lightly-doped drain (LDD) formation, are implant dopants in the S/D regions at low doses of approximately 10^{13} to 10^{14} cm^{-3} [21]. LDD structures are necessary to minimize SCEs such as hot carrier effects. When device dimensions are reduced while bias conditions remain relatively consistent, electric fields increase by definition. High electric fields induce high energy carriers or "hot" carriers that can cause impact ionization. Impact ionization creates additional hole-electron pairs. Carriers with sufficient energy to overcome the 3.2 eV energy barrier between silicon conduction band and the silicon dioxide conduction band can build up in the gate dielectric, which can result in shortening the channel. This shortened channel length is represented as L_{eff} as shown in Figure 5. The fundamental motivation for implementing the LDD structure is to enable a gradual voltage drop over a larger distance between the drain and the channel. Introducing low dose implants in the S/D extension regions enables this gradual voltage drop. Since the degrading effects of high electric fields are

dependent exponentially on electric field, slight decreases can make significant differences to device performance [20].

Halo Implants

Halo implants are also used to form S/D extension regions. Halo implants are high angled implants commonly processed in the same lithography step as the LDD formation. The same types of dopants used in the punch-through-stop implants are used for the halo implant step [18]. Similar to the LDD formation, introducing halo implants in the S/D extension regions helps to create non-uniform doping profiles between the S/D and the channel region to reduce short-channel effects.



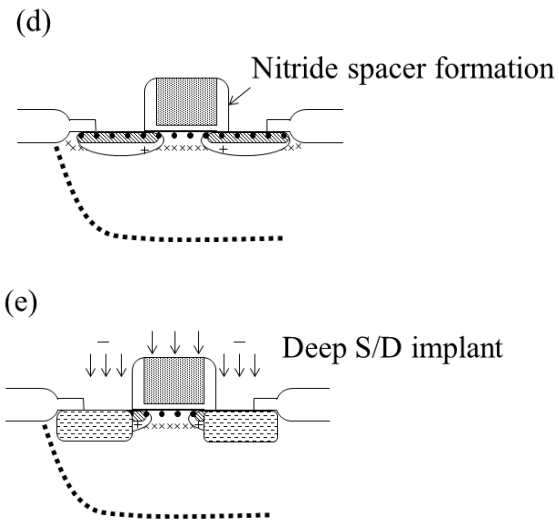


Figure 6. Ion implant and related processing sequences that are connected to the lithography study in this work: (a) Retrograde well, punch-through-stop, and threshold voltage adjustment, (b) S/D extension, (c) halo, (d) nitride spacer formation, and (e) deep S/D ion implantation [18].

The ion implantation that is related to the lithography processing in this paper is a sequential S/D extension and halo implant. These steps are shown as (b) and (c) in Figure 6 as the implant after polysilicon gate patterning but before the nitride spacer formation. It is common to use a high angled implant of 20-30° for the tilt angle of the halo implant [1]. The implant sequence before the S/D extension and halo implant are the retrograde well, punch-through-stop and threshold voltage (V_T) adjustment implant formation, as shown in step (a). The reader is recommended to refer to Plummer *et. al.*[20] for a better understanding of retrograde well and threshold voltage adjustment implants. The nitride spacer is formed after the S/D extension and halo implant, step (c). The block mask discussed in this work is the mask geometries defining the resist opening (or block) for the S/D extension and halo implants.

CHAPTER 3

BACKGROUND

Although block mask lithography processing is a relatively simple step, there are clear device issues related to block mask CD variations. Reflectivity and other light interactions between dense topographies is a key issue to the implant lithography process. In this section, general principles of optical physics will be introduced so that the lithography processing of aerial image formation to Maxwell's equations can be better understood. Maxwell's equations describe the electromagnetic behavior of light and form the basis for lithography simulation calculations. Finally, the simulation structure of Panoramic Technology® will be discussed to understand how the lithography simulation is capable of modeling electromagnetic waves.

3.1 OPTICAL PHYSICS

Projection optics is an integral component of conventional semiconductor lithography. Semiconductor lithograph utilizes a small fraction of the electromagnetic (EM) spectrum. This section will include an introduction to EM spectrum, wave propagation, Snell's law of refraction, basic lenses and wavefronts, interference and diffraction, and the photoelectric effect.

3.1.1 ELECTROMAGNETIC RADIATION

Electromagnetic radiation is the key to many of the current technological advancements. This is because electromagnetic spectrum is a form of energy, given as:

$$E = \frac{hc}{\lambda}, \tag{2}$$

where h is Planck's constant, c is the speed of light in vacuum, and λ is the electromagnetic wavelength. The known electromagnetic spectrum ranges from cosmic rays to radio waves, as shown in Figure 7 [26]. A small portion of this spectrum is visible light, in the wavelength range between violet and red at 400 nm and 760 nm, respectively. Advancements in wireless technology to cloud computing utilize the energies associated with radio waves. Essentially, radio waves are converted into binary codes that are transmitted from a source then received (i.e. a router or a cloud computing

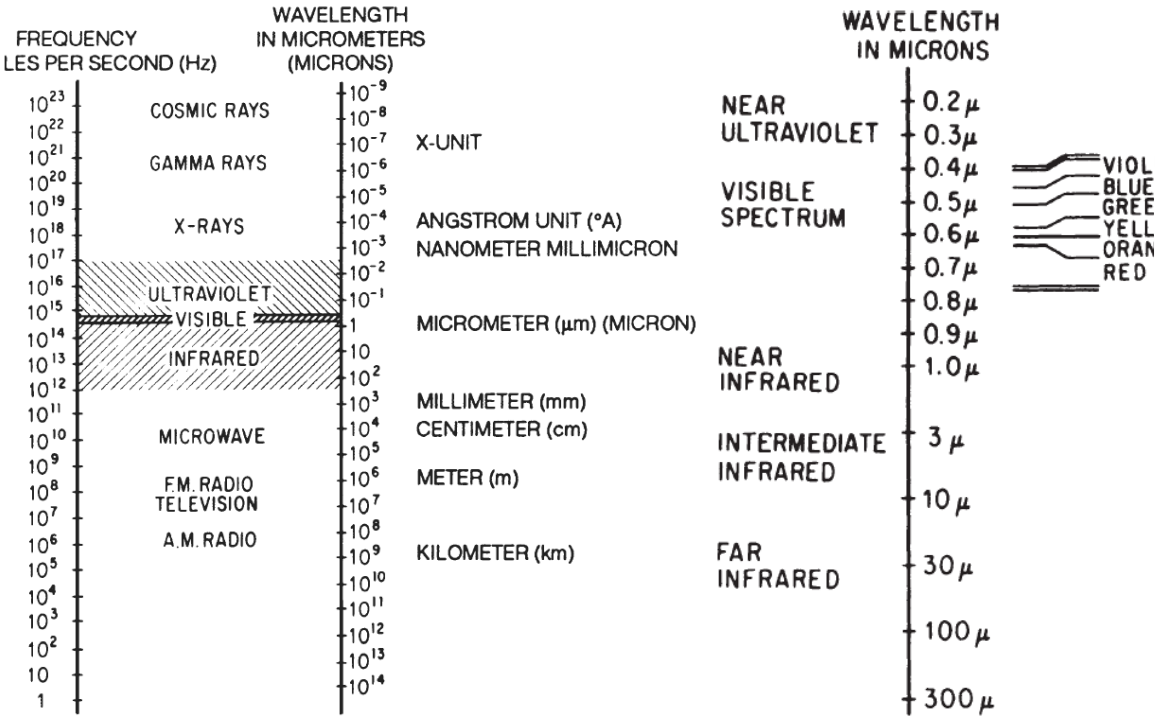


Figure 7. The electromagnetic spectrum shown in a range of frequency (Hz) and wavelength (μm) [26].

control node) to process the given information. Radio waves occur in the kilometer regime. Lithography occurs on the other end of the spectrum.

Lithography technologies benefit from electromagnetic radiation that are in the micro-nanometer regime. The electromagnetic radiation of interest in current lithography is in the intermediate portion of optical to ultraviolet. At this intermediate portion of the spectrum, electromagnetic radiation exhibit both wave and particle-like behavior. EUVL further extends into the nanometer regime that gives rise to even more challenges that are discussed in Section 3.6.D. In lithography, because the wavelength (i.e. 193 nm versus EUV) determines the capabilities of pattern transfer, wave-like behaviors of EM radiation will be explained

3.1.2 LIGHT WAVE PROPAGATION

Wave propagation of light helps to visualize the path a light may take. If the medium is assumed as a vacuum and there exists a point source of light, a spherical wavefront is created as shown in Figure 8. The successive distance of plane waves is the wavelength of radiation. The maximum velocity of light propagation in vacuum is the well-known value of approximately 3×10^{8} m/s. The index of refraction that a light

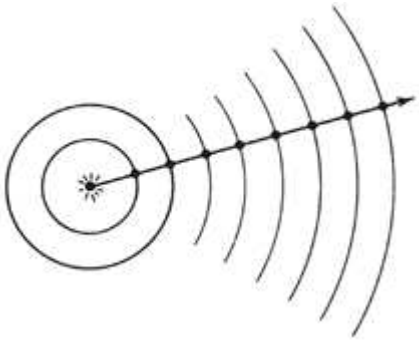


Figure 8. An illustration of a spherical wave from a point source traveling through a vacuum [26].

propagates in a different medium is given as the ratio between the light velocity in vacuum and in the medium. Some of the index of refractions that are commonly used in semiconductor industry is given in Table 2. A ray of light is considered as the path of light that is perpendicular to the normal to the wavefront. If light rays travel through isotropic media, the rays will travel in a straight line until it encounters a boundary change in medium [26]. The ray of light is a convenient representation how light waves propagate.

Table 2. Example values of optical constants (*n* and *k* values) of common processing layers at different source wavelengths [33].

	193		248		365		436	
	n	k	n	k	n	k	n	k
Silicon	0.960	2.88	1.58	3.60	6.41	2.62	4.79	0.175
SiO2	1.56	0.00	1.51	0.00	1.47	0.00	1.47	0.00
Si3N4	2.65	0.180	2.28	0.005	2.11	0.00	2.051	0.00
Aluminum	0.117	2.28	0.190	2.94	0.407	4.43	0.595	5.35
Polysilicon	0.970	2.10	1.69	2.76	3.90	2.66	4.46	1.60
Diazonaphthaquinone (DNQ)/novolac					1.70	0.007	1.67	0.007
PHS CAR			1.76	0.007				

3.1.3 SNELL'S LAW OF REFRACTION

Snell's law explains how light behaves as it travels through different uniform mediums by using geometries of light rays. When there is an object that emits light, the path of the light will travel in a straight path until the light hits an object. This assumes the fact that there are no possible particles in a medium. The presence of particles such as dust or gas particles can result in an energy loss. In this vacuum scenario, light wave propagation can be represented by a plane wavefront instead of spherical because the source is assumed to be infinitely far away. The light vector hitting an object can be thought of as the light hitting a boundary, as shown as the bolded line separating medium

n_1 and n_2 in Figure 9. Snell's law is the mathematical relationship of how a light bends through a boundary between two different mediums:

$$n_1 \sin\theta_1 = n_2 \sin\theta_2 \tag{3}$$

The three possible interactions of an incident light at a medium boundary are refraction, reflection, and/or absorption, and can either be attenuating or non-attenuating. When all of the light is either reflected or refracted or both, the object is considered as non-attenuating. Any other case would be considered as attenuating. If the incident angle (θ_1) and the reflected angle (θ_3) are equal, this is recognized to follow the law of reflection and is considered as a regular or specular reflection. When the light passes through the boundary into the next medium at a different angle, the light is considered to be refracted. A light is always refracted if an incident light travels through two different mediums because n_1 and n_2 are different.

3.1.4 BASIC LENSES AND WAVEFRONTS

A basic visual can now be shown how a wavefront from a point source may travel through a lens element. First, assume a biconvex lens that has a higher index of refraction than that of the medium the wavefront is traveling through as shown in Figure

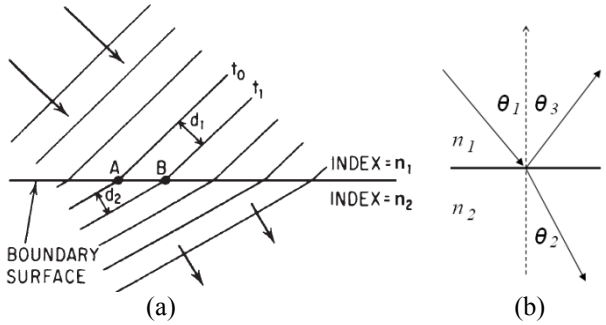


Figure 9. An illustration of Snell's law showing how a (a) parallel wavefront [26] can also be represented by (b) light vectors traveling through mediums n_1 and n_2.

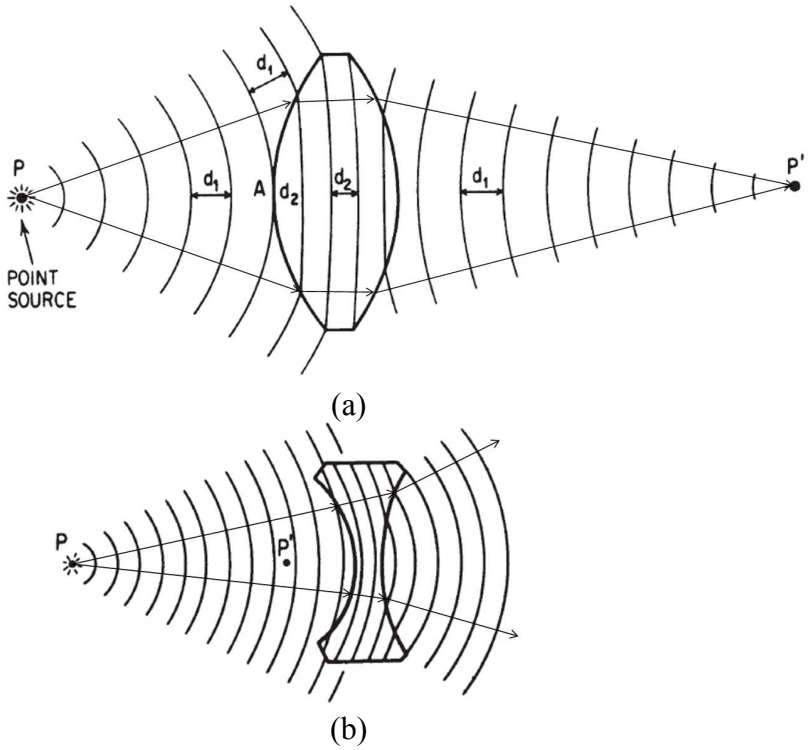


Figure 10. A schematic illustration showing the interaction of spherical wavefronts with its corresponding light rays traveling in vacuum with a (a) positive, convex and (b) negative, concave lens element with a higher index of refraction than 1.0 [26].

10(a). As the wavefront approaches the lens boundary, there is a retardation effect inside the lens and the interval d_2 is less than that of d_1. The retardation effect is greater at thicker regions in the lens. In fact, the exiting wavefronts inside the lens are thick enough to eventually change the curvature of the wave. The lens effectively changed a diverging point source wavefront at point *P* to a converging light wave at point *P'*. For this reason, the type of lens is called a converging or positive lens. A similar analysis is made for a biconcave lens. The wavefronts are shown Figure 9(b). However, the thicker regions at the edge and the central concavity of the lens increase the divergence of the wavefront. Unlike the convex lens, there is no place for point *P'* to image a concentration of light with the given imaging system. For the concave lens, the image point *P'* occurs prior to

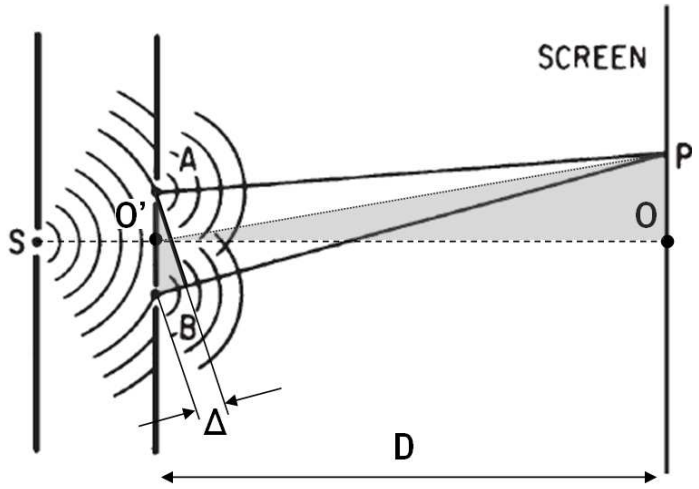


Figure 11. Young's diffraction experiment with geometries [26].

the lens, creating a virtual image. The possible light rays path in Figure 9 shows the relationship between how a light wave/ray travels through a positive lens element [26].

3.1.5 INTERFERENCE AND DIFFRACTION

In most imaging systems, since there are several wavefronts, interference and diffraction are important phenomena to understand. Young's experiment is the famous demonstration of both diffraction and interference. Just like two ripples of water coming together, when two waves (e.g. crests) come together, they add for constructive interference, while a crest and trough result in a cancellation for destructive interference. The setup of the experiment is shown as Figure 11.

In Young's diffraction experiment, there is a source point at *S* and at a distance away, there are two slits *A* and *B* in an opaque cover and at a distance *D*, there is a screen to image the interference and diffraction effects. Constructive interference is represented as an illuminated spot at point *P* on the screen because the two wavefronts that arrive simultaneously are in phase. Destructive interference is represented by a dark spot at point *P* because the two waves are out of phase. Assuming that slits *A* and *B* are

equidistant to *S*, it was found that when the distance of *AP* and *BP* were exactly equal, or if they differed exactly by an integral number of wavelengths, the image was a light spot at *P*. When *AP* and *BP* differed by one-half wavelength, the image was dark at point *P*. This experimentally showed the wave-like properties of light [26].

If a monochromatic light source is used, Young's diffraction experiment can determine the wavelength of the illuminating source. A monochromatic light is a single wavelength light. The relative distance between the slits to the screen and the slit separation should be at least about three orders of magnitude in difference so that small angle trigonometric approximations can be made. In Figure 11, by basic geometry, the shaded regions are similar triangles. Assuming small angles, $D \approx O'P$. This allows the following relationships:

$$\Delta = AB\sin\theta = \frac{AB \cdot OP}{D},$$

where Δ is the path difference between the two wavefronts. The above equation can be rearranged to give the following equation:

$$OP = \frac{\Delta \cdot D}{AB}. \tag{4}$$

Since OP is the distance along the screen, and the path difference is related to the monochromatic wavelength, the distance of dark and bright spots can be calculated as:

$$OP(dark) = \frac{\pm\left(m + 1/2\right)\lambda \cdot D}{AB} \tag{5}$$

$$OP(bright) = \frac{\pm m\lambda \cdot D}{AB}, \tag{6}$$

where *m* is an integer value corresponding to the order of interference.

Young's experiment can also be applied to thin-film interference with additional phase difference. The path difference that determines the phase relationship is shown as AB and BC in Figure 12. The Δ value in this case is $n\left(AB + BS\right)$, where *n* is the index of

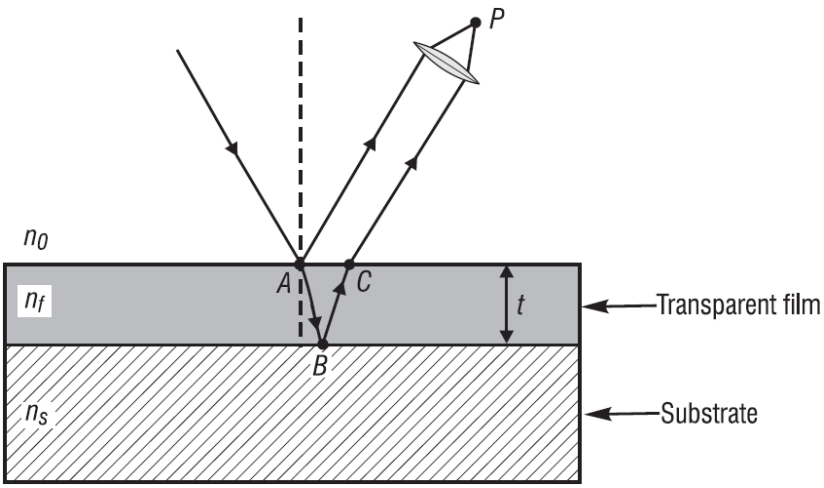


Figure 12. Light interaction with a transparent film with a propagation path between three different index films [27].

refraction. Since there are orders of magnitude difference between film thickness and the path length of light rays from the source to detector, an approximation can again be made to give a Δ value of:

$$\Delta = n(2t). \tag{7}$$

However, there are established facts that a light wave traveling from a medium of lower index of refraction to a higher medium automatically undergoes a phase change of 180° upon reflection. There is no phase change upon reflection when a light wave travels from higher to lower index of refraction. Also, the wavelength of light in a refractive index of n is given as $\lambda_n = \lambda_0 / n$, where λ_0 is the wavelength for vacuum (or approximately air) [27]. If the light propagates through the three different media relative to the previous as low to high to low, the light reflection at point *A* has a phase shift of 180°. This requires an additional $\lambda_0 / 2$ phase shift when light reflects at point *A*, leading to a destructive interference. The optical path difference can be summarized by the following equation:

$$\Delta_p + \Delta_r = m\lambda,\ (m = 1,2,3,...), \tag{8}$$

where Δ_p is optical path difference of the film and Δ_r is the equivalent path difference introduced at the reflections at points A and B. The optical path difference of the film can be determined by Equation 7. The constructive interference at normal incidence is:

$$2n_f t + \frac{\lambda_0}{2} = m\lambda_0,\ (m = 1,2,3,...). \tag{9}$$

The destructive interference at normal incidence is:

$$2n_f t + \frac{\lambda_0}{2} = (m + 1/2)\lambda_0,\ (m = 1,2,3,...). \tag{10}$$

Phase change can be defined by the perpendicular and parallel component of the reflection irradiance. For a source wavelength that is far away (relative to the wafer geometries), the perpendicular component of reflectivity is of interest.

3.2 IMAGE FORMATION

The fundamental goal for lithography processing is to accurately form in a resist the projected image of a mask which represents a design layer of a semiconductor chip. This process of pattern transfer requires an understanding of the concepts of aerial images, photoresist imaging, and exposure kinetics of the chemistry. Each of these will be discussed in the following sections.

3.2.1 AERIAL IMAGE

Aerial image is light distribution of an image plane based on the physical properties of light wave propagation. This image plane is a distribution prior to any interaction between light and resist and/or multiple-reflecting surfaces. The simplistic

basis of light wave propagation is its diffraction-limited nature. The aperture of the imaging lens is the only limiting factor on resolution. Since engineering an optimal lens design is based on spherical wavefront, it is a key to understanding diffraction-limited image formation.

3.2.2 SPHERICAL WAVEFRONTS

Mathematical relationships are used to determine the effect of a finite numerical aperture on spherical wavefronts [28]. A basic three-dimensional setup for a spherical wavefront is shown in Figure 13. Assuming a perfect spherical wavefront, an electric field *E*, of the light wave is given as:

$$E = Ae^{ikr}, \tag{11}$$

where A is the magnitude of the electric field, r is the radius distance between the image point P and the spherical wavefront, and k is defined as the wave number:

$$k \equiv 2n\pi / \lambda_0 . \tag{12}$$

Only electric field is used in this section to represent the electromagnetic behavior of light because light intensities for photoresists are proportional to $|E|^2$ [28]. Due to its converging nature of light propagation, spherical wavefront produces the sharpest image. However, spherical wavefronts are constrained by the finite angle θ capturing the light source, or otherwise known as numerical aperture (NA). NA is given as:

$$NA = n \sin\theta . \tag{13}$$

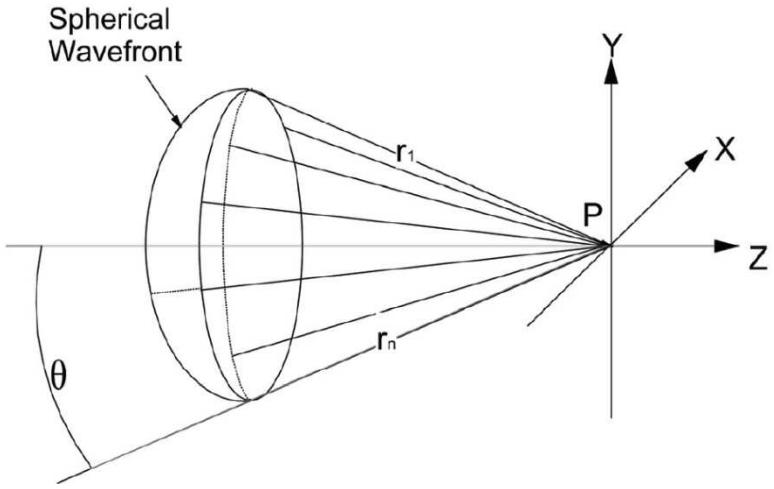


Figure 13. A three-dimensional schematic illustration of a spherical wavefront showing multiple radius lengths r to r_n, between the image point P and a point at the wavefront [28].

The effect of a finite NA on intensity from a spherical wavefront is analyzed by referring to the illustration shown in Figure 14. Let θ_A be the angle defining NA, R be an arbitrary point on the spherical wavefront at an angle θ with respect to the optical axis z. Since $P(x,0,0)$ has no y dependence, the electric field of the image will be dependent on integrating all angles confined in the NA angle θ_A. The integral is given as:

$$E_x\left[P(x,0,0)\right] = \frac{1}{2\theta_A}\int_{-\theta_A}^{\theta_A} Ae^{iks}\cos\theta d\theta\,, \tag{14}$$

where s is the distance between R and P. The cos θ term is included to take the z-component of light propagation. Since a spherical wavefront converges to a common point, x can be assumed to be close to 0. By small angle approximation, $R_0P = RP_0 = R_0P_0$. Again, by geometry and trigonometry, distance d is given as $x\sin\theta$. The relationship between r, which is simply the radius or R_0P_0, and s is just $r - x\sin\theta$. Plugging the s value and completing the integral gives the two equations:

$$E_x\left[P(x,0,0)\right] = \frac{1}{2\theta_A}\int_{-\theta_A}^{\theta_A} Ae^{ik(r-x\sin\theta)}\cos\theta d\theta\,, \tag{15}$$

$$E_x[P(x,0,0)] = \frac{Ae^{ikr}\sin(kx\sin\theta_A)}{ikx\sin\theta_A}. \tag{16}$$

Now substituting for NA from Equation 13 gives:

$$E_x[P(x,0,0)] = \frac{Ae^{ikr}\sin\left(2\pi\frac{NA}{\lambda}x\right)}{2\pi\frac{NA}{\lambda}x}. \tag{17}$$

Equation 17 indicates that the intensity of the electric field is not zero at the image point P_0, as expected. In fact, the highest intensity should be expected at P_0. However, the sine function allows multiple instances where the intensity can equal zero. The first instance is when:

$$2\pi\frac{NA}{\lambda}x = \pi .$$

Solving for x gives:

$$x = 0.5\frac{\lambda}{NA}. \tag{18}$$

The coefficient factor of 0.5 is related to the k_1 in the Rayleigh CD criterion in Equation 1. It is interesting to note that even with a perfect spherical case, there are limits to the image spot size confined by wavelength and NA [28].

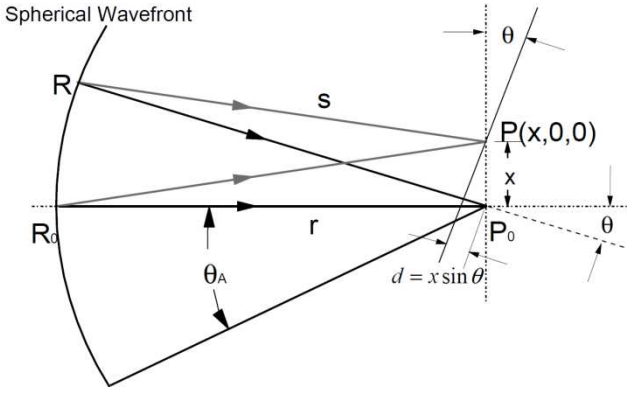


Figure 14. A two-dimensional illustration of the spherical wavefront for calculation of the effect of a finite NA on $E[P(x,0,0)]$ [28].

In reality, a point image from a spherical wavefront has deviations caused aberrations, making the image complicated to analyze because it is no longer diffraction limited. The aberration is characterized by two major coefficients – Seidel and Zernike aberration coefficients.

3.2.3 PHOTORESIST IMAGING

The wave-like behavior of light helps to characterize its possible paths. The focus will now shift towards how light interacts with materials. Light interaction is based on its electromagnetic wave characteristics. A combination of standing waves, swing curves, exposure kinetics and chemical amplifications allow the conversion of aerial images into a spatial distribution in the photoresist.

Photoresist

Photochemical reaction is chemistry used for light exposure. Photoresist is a photosensitive formulation that when developed forms a three-dimensional image profile based on light exposure projected from a mask pattern. The primary function of a photoresist is to act as a protection mask for selective dry/wet etching and ion-stopping. There are essentially two classifications of photoresists – positive and negative. Positive photoresist becomes more soluble in exposed areas. Negative photoresist becomes less soluble in exposed areas. The addition of a photoresist layer also adds as another medium for light waves to propagate and interact.

In general, photoresists have four main components:

1. base resin
2. photoactive compound (PAC)

3. solvent
4. quencher

Each of these four components has a specific function that allows a resist to chemically alter its structure to imprint a pattern. A base resin determines the mechanical and thermal properties and may have a component that reacts with products of the photochemical reaction. A common base resin used for *i*-line (365nm) positive resists is phenolic novolak. Phenolic string formation is seen as the principal interaction between inhibitors and the novolak resin. Phenolic strings survive the film casting process because of its strong chains of ionic bonds with several hydroxyl OH groups. There is also an interaction between acceptors and phenols. Due to their length and average acidity of their OH groups, phenolic strings have salient properties. The protons of the phenolic strings are less likely to dissociate and the acidity of the phenols will be lower than that of free phenols [29]. The PAC determines the sensitivity to radiation. Diazonaphthoquinine (DNQ) is a common PAC inhibitor. The solvent component controls the viscosity and allows certain areas of photoresist to be developed. An example of a photochemical reaction is shown in Figure 16. The limit to using certain photoresists is dependent on the sensitivity to the wavelength of exposure.

In the deep UV range of 248 and 193 nm, chemical amplification resists (CAR) are used for higher sensitivity at that range of wavelength. CAR differs from a conventional photoresist in that it is a two-step chemical process. Because of its sensitivity to light, photoacid generator (PAG) molecules decompose after light exposure. The reaction products then create more photoproducts that help to catalyze the decomposing process. A single photon can lead to 1000 decomposition reactions [30].

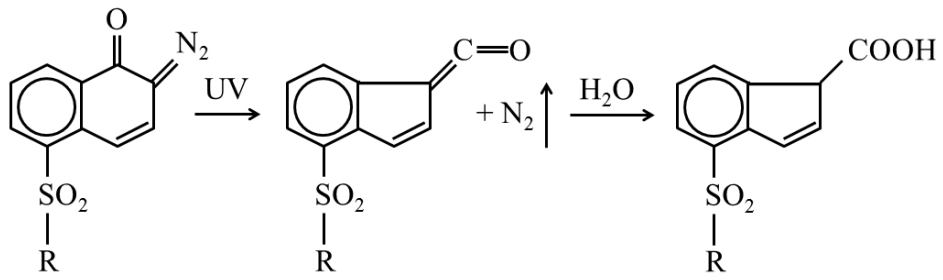


Figure 16. Diazonaphthoquinine (DNQ)-novolak resist chemical reaction [30].

Photoacid generators that are used commonly to form acid formation to promote chain reactions are based on onium salts. Some examples of onium salts are shown in Figure 15. When the onium salt undergo a photogenerated reaction (salt→ acid) the salt dissociates and forms the acid. The disassociation of the salt after exposure is due to the fact that "x-group" (e.g. AsF_6 for Diphenyliodonium salt) attached is bonded by ionic bonds [33].

A t-BOC polystyrene (TBS) based resist chemical reaction is shown in Figure 17 to show a 248 nm chemical amplification positive resist system. Prior to a post-exposure bake (PEB), the poly(t-butoxycarbonyloxystyrene) is a protected hydrophobic polymer.

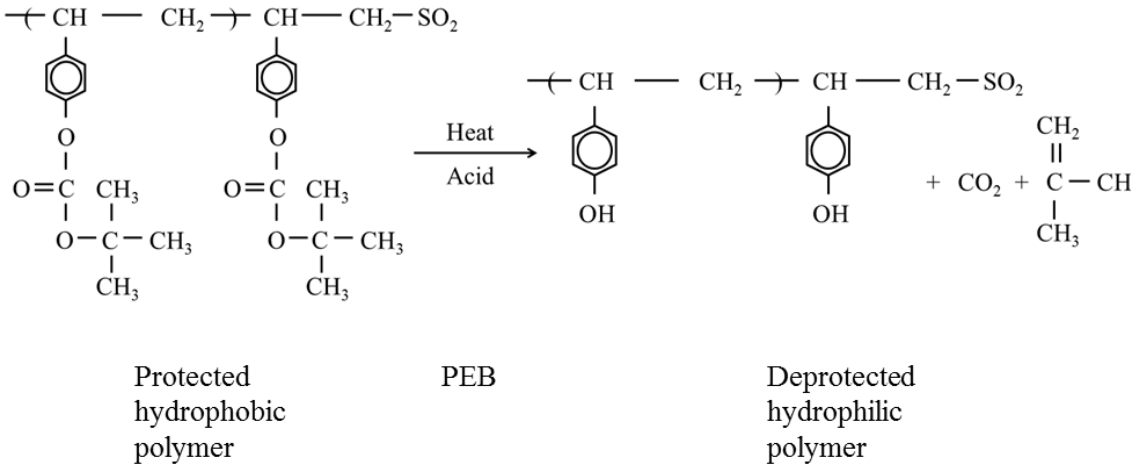


Figure 17. 248 nm positive resist chemical reaction using a PAG [33].

An acid generator of onium salt cationic photoinitiator is introduced to the polymer. After exposure, the salt component, or the "x-group," disassociates and forms an acid. Once the acid is introduced to the polymer and PEB occurs, the reaction changes the system to a deprotected hydrophilic polymer. The exposed regions cause these hydrophilic polymers to be developed away. The fourth component of a formulation is typically a base, also known as a quencher, which helps prevent unwanted parasitic reactions.

Swing Curves and Standing Waves

If there are multiple media with different refractive indices, an incident light will be refracted, reflected, and transmitted, giving rise to standing waves and swing curves. Both of these effects seen in photoresists are detrimental to the integrity of the resist profile. A swing curve refers to the lithographic parameters such as resist linewidth

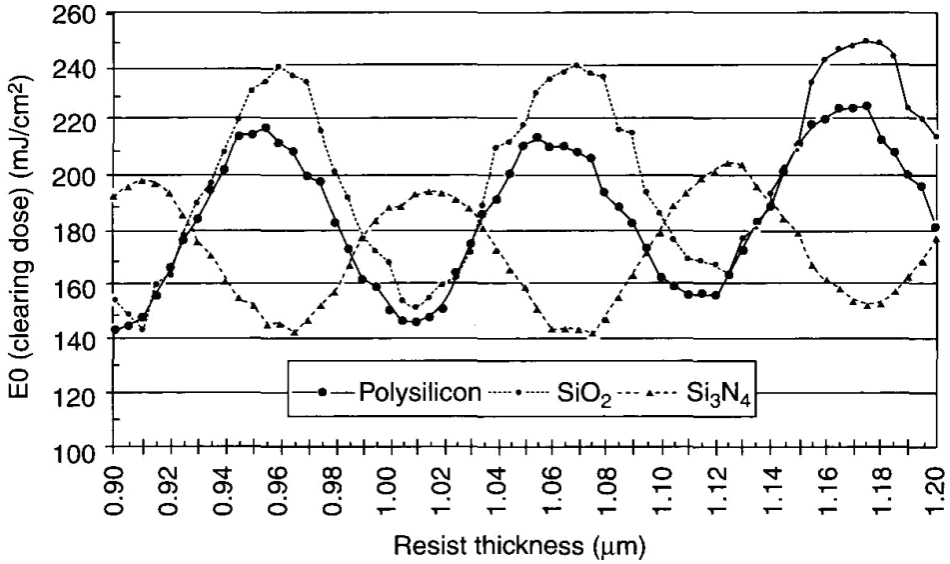


Figure 18. A swing curve of clearing dose (E_0) with varying 365 nm wavelength resist thickness over polysilicon, oxide (3700 Å), and nitride film (1260 Å) [31].

(CD), dose-to-clear, and reflectivity, that are sensitive to thicknesses of film stack layers. If the maximum to minimum swing is large, the processing parameters become sensitive to thicknesses of film stack layers. An example swing curve of clearing dose E_0, is shown in Figure 18. E_0 swing curves can help determine optimum resist thickness. Analyzing the graph, a minimum dose to clear corresponds to a constructive interference and maximum intensity at the resist base. Destructive interference leads to a maximum E_0 requirement. At midnodal positions, small deviations in the resist thickness correspond to large difference in dose requirements, which lead to the least amount of coating process latitude. The optimum resist thickness is when there is the least change in clearing dose, which occurs at a minimum or maximum. The minimum clearing dose occurs is not the desired thickness because the direction of change in E_0 is toward higher clearing dose values. This indicates that unwanted scumming of resist features from underexposure will occur. The optimum resist thickness occurs at a maximum peak because a small resist thickness change results in a small decrease in dose requirement at a maximum clearing dose. At this resist thickness, the clearing dose has a slightly lower throughput compared to the minimum clearing dose but does not result in scumming [31].

Standing waves occur when two waves traveling in opposite directions combine. This can occur in lithography because light propagating through a resist can reflect off the substrate or any underlying layers. If the reflected light significantly interferes with the incoming wave, a standing wave pattern is seen on the resist sidewall profile.

Absorption in Exposure

When a photoresist is exposed to a light source, there is absorption of radiation that result in photochemical induced changes and subsequently change dissolution properties. Absorption characteristics of photoresist influence its resolution and process capabilities. Maximum transmission (the ability for highest light intensity to reach the bottom of the resist) and absorption with highest sensitivity are desired. Beer's law explains that intensity has an exponential loss of absorbance in the bulk of a material with the following relationship:

$$I(z) = I_0 e^{-\alpha z}, \tag{19}$$

where *I(z)* is the intensity of light traveling in the z-direction through a medium, α is the absorption coefficient, z is the distance the light traveled through the medium and I_0 is the intensity at $z = 0$. Since transmission is related to intensity with the ratio I/I_0, Beer's law ultimately relates transmission and absorbance [31]. The absorption coefficient α, is given as:

$$\alpha = 4\pi\kappa / \lambda, \tag{20}$$

where κ is the imaginary part of the index of refraction.

A mathematical representation of the absorption mechanics can be made with the basic understanding of photoresist chemical reactions. A typical photoresist used in 436 and 365 nm lithography has three components of resin *R*, sensitizer *M*, and solvent *S*. As chemical reactions are consumed, products *P* are generated to complete a reaction. The first exposure reaction is:

$$M \xrightarrow{UV} P$$

Each component has an associated absorption coefficient. Using Beer's law,

$$\alpha = \alpha_M M + \alpha_P P + \alpha_R R + \alpha_S S \tag{21}$$

The products being produced can be mathematically written as:

$$P = M_0 - M\,, \tag{22}$$

where M_0 is the sensitizer without any exposure. Since the absorption coefficient is changing during exposure, Equation 21 can be rewritten as:

$$\alpha = Am + B\,, \tag{23}$$

where A and B are the first two absorption parameters known as the Dill parameters. The exposure dependent absorption of the resist (parameter A) is given as:

$$A = (\alpha_M - \alpha_P)M_0\,; \tag{24}$$

The exposure independent absorption (parameter B) is given as:

$$B = \alpha_P M_0 + \alpha_R R + \alpha_S S\,; \tag{25}$$

The relative sensitizer concentration is given as:

$$m = M / M_0\,. \tag{26}$$

Graphically, this can be shown as in Figure 19 [32]. If the A/B parameter is greater, it indicates more efficiency for absorption because there is more of the photoproduct reaction occurring.

3.3 ACCURACY OF PATTERN TRANSFER

The accuracy of pattern transfer is based on the ability to control the CD and overlay specifications effectively. Some examples of techniques used for lithographic control are BARC layer processing and utilizing resolution enhancement techniques (RET).

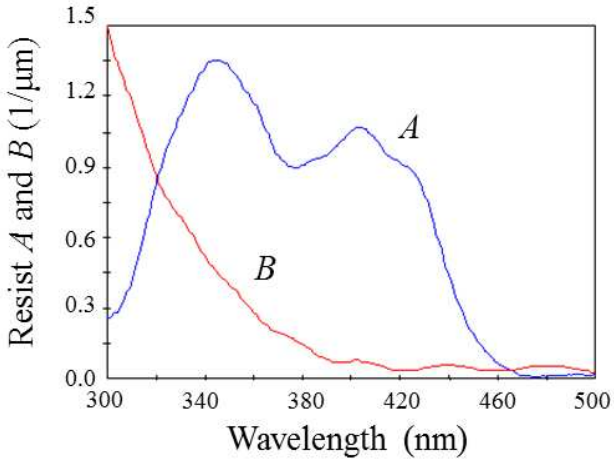


Figure 19. A graphical representation of *A* and *B* resist parameters as a function of wavelength in the deep UV range for a typical 436 nm resist [32].

CD Control

The two main reasons for CD control are to limit electrical device degradation affected by printed geometries and to minimize the variability of lithography processing. An effective gate length is directly related to the transit time of a device. Smaller gate lengths can also lead to SCE because threshold voltage is dependent on gate length geometries. A tighter distribution of gate CD can ultimately reduce failing devices resulting from leakage current or devices that are too slow.

Overlay Control

While geometry definition is important, the accuracy of placing a defined pattern on top of the underlying layer is just as critical. Overlay errors can increase device leakage or even short an electrical connection. Electrical shortage would require a drastic error because most features are designed to take into account expected errors in overlay. Better overlay control leads to an ability to shrink designs, decrease die size, increase die

per wafer, and ultimately more devices per wafer. Current overlay measurement techniques involve direct optical measurement of displacement using standardized structures such as “box-in-box” method.

Reflectivity Control - BARC

When light propagates through multiple media with different indices, each interface gives rise to reflection that needs to be minimized in the photoresist. Engineering a bottom anti-reflection coating (BARC) layer can help to optimize the reflectivity issues in the photoresist. Minimizing standing waves and swing curves requires three tasks.

1. Optimize the BARC
2. Optimize the resist thickness
3. Understand the sensitivity to BARC, resist, and film stack variations

In consideration to a lithography process, all of the light rays, from reflected to

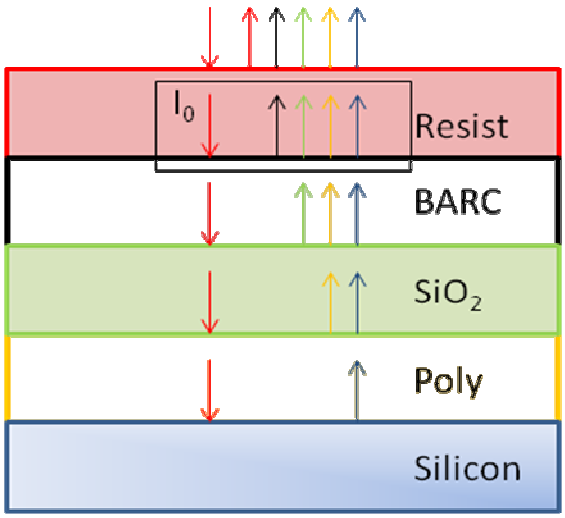


Figure 20. A geometrical representation of a film stack and incident light source irradiance I_0 is shown. The resulting total irradiance at the resist and vacuum interface is the total summation of any light refracted and reflected from underlying layers.

refracted to absorbed, need to be considered when analyzing a total irradiance. Instead of considering only a two medium system, each different environment a light interacts with needs to be considered. A simple geometrical representation of light irradiance is shown in Figure 20, where the light incidence is normal to the surface. If the initial irradiance I_0 is assumed to initially travel through vacuum, the different mediums that the light refracted and reflected are through the medium or thin films of resist, silicon dioxide, and silicon substrate. The total amount of irradiance between any interfaces can be calculated if the initial irradiance, thicknesses, and optical constants of the film stack are known.

For example, using the illustration in Figure 20 with the various thin films, a calculation of total irradiance of interest can be made by the equation:

$$R = \frac{I_{total}}{I_0}.$$

The reflection at normal incidence is given by the equation:

$$R_{\perp} = \left| \frac{n_1^* - n_2^*}{n_1^* + n_2^*} \right|^2,$$

Where the n^* is real component of index of refraction and the subscript 1 and 2 refers to the different processing layers. The total irradiance can then be estimated by the different irradiances derived from Beer's law (Equation 19) and its associated phase change (φ) with the simplified equation:

$$I_{total} = \left(\sqrt{I_1} \cos\varphi_1 + \sqrt{I_2} \cos\varphi_2 + \sqrt{I_3} \cos\varphi_3 + \sqrt{I_4} \cos\varphi_4 \right)^2. \quad (27)$$

Equation 27 can be summarized that the total irradiance can be approximated by the square of the sum of the electric fields [33]. With known optical constants and material thicknesses, substrate reflectivity for an exposure system can be simulated.

Optical Proximity Correction

Resolution can further be improved by resolution enhancement technologies (RETs) such as optical proximity correction (OPC). Improving the design of an imaging system has been called wavefront engineering by Marc Levenson [34]. OPC is necessary because of proximity effects that give rise to variations to linewidth features because of optical bias caused by nearby features. Since the CD variation is considered as a systematic error, rule based proximity correction can be implemented to compensate for the difference.

3.4 MATHEMATICAL DESCRIPTION OF LIGHT

Before Panoramic Technology ® simulation software is introduced, Maxwell's equations should be discussed because of its important to the mathematical description of light. The calculation made in the modeling simulator uses the finite-difference time-domain (FDTD) method applied to Maxwell's equations. In general, Maxwell's equations are the fundamental equations that are related to the imaging process in the lithography industry [32]. The differential form of Maxwell's Equations with a magnetic and/or polarizable media is given by:

$$\vec{\nabla} \times \vec{E} = -\frac{\partial \vec{B}}{\partial t} \tag{M.1}$$

$$\vec{\nabla} \times \vec{H} = \vec{J} + \frac{\partial \vec{D}}{\partial t} \tag{M.2}$$

$$\vec{\nabla} \circ \vec{D} = \rho \tag{M.3}$$

$$\vec{\nabla} \circ \vec{B} = 0 \tag{M.4}$$

For an in-depth analysis of Maxwell's equations, the author recommends to refer to Fleisch [37].

3.4.1 GAUSS' LAW FOR ELECTRIC FIELDS

The general idea of Gauss' law for electric fields is that the electric flux through any closed surface is proportional to the total enclosed charge. The integral form of Gauss' law is given as:

$$\oint_S \vec{E} \circ \hat{n} da = \frac{q_{enc}}{\varepsilon_0}, \tag{M.5}$$

where $\vec{E}$ is the electric field, $\hat{n}$ is the unit vector normal to a given surface, *da* is an increment of surface area, q_{enc} is the charge enclosed, and ε_0 is the electric permittivity of the free space.

An electric field is an electric force surrounding an object that has electrical properties. Mathematically, electric field is a vector quantity defined as the electrical force vector divided by a small test charge. Depending on the kind of enclosed charge (*i.e.*, point charge, conducting sphere, infinite line charge, infinite flat plane, etc), there are well known equations to calculate electric fields. For example, the simplest case of a point charge *q*, with respect to radius *r*, is given as:

$$\vec{E} = \frac{1}{4\pi\varepsilon_0} \frac{q}{r^2} \hat{r}. \tag{M.6}$$

As a convention, electric field lines originate from a positive charge and terminate on a negative charge. A positive charge can be thought of as a source (field lines diverging) and a negative charge as a sink (field lines converging).

The task of analyzing the left side of Gauss' law in integral form (Equation M.5) is to determine the net flux of the electric field. The three main tasks are to define a closed surface, determine the electric field, and calculate the integral of the electric flux with respect to the perpendicular orientation of the enclosed area. Determining the shape for defining the surface area surrounding the charge of interest helps to simplify calculations. The dot product operation indicates that any field lines parallel the surface are zero, while perpendicular lines are a magnitude of one, as an example, using a circular enclosed area for a point charge. Once the simplest shape is defined, the integral can be calculated to obtain the electric flux.

The right side of Gauss' law in integral form only involves the *net* charge enclosed which determines the amount of flux produced. Any electric flux that originates and terminates outside the defined surface has a net flux of zero. If there are several charges, the total enclosed charge is simply the summation of all the individual charges. More interesting situations are when there are charged objects that are contained throughout a line, surface or volume. The total charge can be determined by taking its integral if there is a known charge density and how it varies along a line, surface or volume.

The important distinction between Gauss' law in integral form (Equation M.5) and the differential form (Equation M.3) is that the differential form gives a mathematical representation of divergence of an electric field and the charge density at individual points in space. The integral form determines the electric field normal over a closed surface. The right hand side of Equation M.3 is simple to understand. It is the charge density associated at the location with the diverging electric field. The idea of divergence

and convergence can be thought of how vector fields "flow" in the area of interest. Convergence is used to describe the rate at which electric field lines "flow" towards a negative charge. Conversely, divergence is the rate of change of electric field lines away from a positive charge. Therefore, a positive divergence corresponds to a "source whereas a negative divergence represents a "sink."

3.4.2 GAUSS' LAW FOR MAGNETIC FIELDS

A charged particle can also give rise to a magnetic field. An important distinction is that a magnetic field is caused by a *moving* charged particle. The magnetic force of the particle moving at a velocity *v*, with respect to a magnetic field *B*, is give by the Lorentz equation:

$$\vec{F}_B = q\vec{v} \times \vec{B}\,. \tag{M.7}$$

The definition of current is a moving charge, which means that it has a magnetic field. Biot-Savart law gives a mathematical relationship between magnetic field and current:

$$d\vec{B} = \frac{\mu_0}{4\pi}\frac{Id\vec{l} \times \hat{r}}{r^2}, \tag{M.8}$$

where μ_0 is the permeability of free space and the radius (*r*) is the distance away from the incremental length vector (*dl*). Simple objects such as infinite or segmented straight wire, circular loop, solenoid with *N* turns, and torus with *N* turns have equations for currents that can be easily referenced [37].

The fundamental idea behind Gauss' law for magnetic fields is that the total magnetic flux through any closed surface is zero:

$$\oint_S B \circ \hat{n} da = 0\,. \tag{M.9}$$

While both Gauss' law for electric and magnetic fields deal with field lines, the main difference is that magnetic fields occurs in pairs between "north" and "south" poles. Because of the reasoning described in the previous section, Equation M.9 must equal zero.

The differential form of Gauss' law for magnetic field as shown in Equation M.4, states that the divergence of a magnetic field is always zero. The value of magnetic charge density is zero because the divergence is the rate of change of individual points in space and there is no such thing as an isolated magnetic pole. Magnetic field lines are solenoidal so magnetic fields cannot be considered as point sources or sinks as described by electric fields, but are rather closed loops with no start or end. Essentially, Equations M.4 and M.9 indicate that magnetic monopoles do not exist.

3.4.3 FARADAY'S LAW OF INDUCTION

Faraday's law of induction states that a changing magnetic flux through a surface induces an electric field around the closed loop of that boundary:

$$\oint_C \vec{E} \circ d\vec{l} = -\frac{d}{dt}\int_S \vec{B} \circ \hat{n} da \,. \tag{M.10}$$

The key purpose of Farday's law is to demonstrate the presence of a *changing* magnetic flux that produces an electric field in the loop enclosed around the flux. Therefore, if a bar magnet is moving inside a loop of wire (or the loop of wire is moving relative to the bar magnet), a loop of wire will carry a circulating electric field.

The left side of Equation M.10 is known as the electromotive force, or emf. An emf is a standard terminology for sources of electrical energy with a unit of volts. By definition, an electric field is the force per unit charge, which is capable of forcing

charged particles in motion within a wire. The electric field produced by the change in magnetic field induces an electric current and therefore generating the induced emf. The work required to move a unit charge for a given distance is given by:

$$\oint_C \vec{E} \circ d\vec{l} = \oint_C \frac{\vec{F}}{q} \circ d\vec{l} = \frac{\oint_C \vec{F} \circ d\vec{l}}{q} = \frac{W}{q}.$$

The right side of Equation 10 is simply the rate of change of magnetic flux. One of the critical differences between Gauss' law for magnetic fields and Faraday's law is the integral over *any* surface *S* and not a *closed* surface. This means that the rate of change of magnetic flux is not necessarily zero. Some possibilities of changing magnetic flux are the magnitude of magnetic field may change; the angle between the surface normal and the magnetic field may vary; and the size of surface may change while keeping the magnetic field and the surface normal constant. The negative sign is known as Lenz's law which indicates the direction of the induced electric field. The induced circular electric field produces its own magnetic field. If a bar magnet is moving closer through a wire loop, then there is an increase in the magnetic flux that induces an electric field. Lenz's law indicates that the induced electric field flows in a direction that opposes the magnetic field from the bar magnet. If a bar magnetic is moving away from the wire loop, there is less magnetic flux present. Therefore the direction of the induced electric field flows in the opposite direction to compensate for the loss of bar magnetic field.

The differential form of Faraday's law states that a change in magnetic field produces a circulating electric field, mathematically represented as:

$$\vec{\nabla} \times \vec{E} = -\frac{\partial \vec{B}}{\partial t}. \qquad \text{(M.11)}$$

The del operator and the cross product with a vector describes the nature of the curling motion of a field, so Equation M.11 validates the previous explanation of a changing magnetic field which induces a circulating electric field. Again, the presence of a negative sign indicates the opposing direction of the induced electric field to its magnetic field. The usefulness of the differential form of Faraday's law is shown below when applied to the wave equation.

3.4.4 AMPERE-MAXWELL LAW

The Ampere-Maxwell law in integral form is given as:

$$\oint_C \vec{B} \circ d\vec{l} = \mu_0 \left(I_{enc} + \varepsilon_0 \frac{d}{dt} \int_S \vec{E} \circ \hat{n} da \right). \tag{M.12}$$

This equation is considered as a joint contribution between Ampere and Maxwell. Ampere's law related a steady electric current to a circulating magnetic field; however, Ampere's law only applied to static situations with steady currents. Maxwell made the integral discovery that a changing electric flux was an important aspect for a comprehensive understanding of the theory of electromagnetism. The combination of both their contributions develops a closer representation of actual relationship between magnetic field and electric flux.

Splitting Equation M.12 in three parts helps to understand the main idea behind Ampere-Maxwell law. Similar to Faraday's law of induction (Equation M.10), the left side describes the magnetic field around closed loop path *C*. This is the outcome from the two sources on the right side of the equation. These sources are a steady conduction current and a changing electric flux through any surface *S* bound by path *C*. Therefore, a circulating magnetic field is produced along a closed path if the enclosed surface has an

electric current and/or changing electric flux. In contrast to Faraday's law being a magnetic induction, Ampere-Maxwell's law can be thought of as an electric induction. However, the magnetic field induced by the changing electric flux is extremely weak, which made its existence hard to detect in the nineteenth century.

The differential form of Ampere-Maxwell's law shown in Equation M.2 explains an existence of the curl of a magnetic field produced from both an electric current and an electric field changing with time. Like the integral form, there are two source terms that produce the magnetic field. The current density is a vector term in a direction relative to the unit normal to the cross-sectional area of the direction of current. The second term on the right side of Equation M.2 involves the rate of change of the electric field with time. This factor is well-understood to be the displacement current because of its effect as the physical displacement of electrical particles caused by the elastic deformation of magnetic vortices. Although the units are consistent with a current density term, there is no physical current flowing or any charges present. The main concept to understand is that this change in electric field induces the curl of a magnetic field. This time-dependent term contributed to the comprehensive theory for electromagnetism.

3.4.5 TEMPEST AND FINITE-DIFFERENCE TIME-DOMAIN (FDTD)

Time-domain Electromagnetic Massively Parallel Evaluation of Scattering from Topography (TEMPEST) is an efficient technique to study complex wafer topographies to understand study issues of alignment, metrology and research in lithography. TEMPEST, first implemented by R. Guerrieri *et al* [38], uses the Finite-Difference Time-Domain (FDTD) algorithm introduced by Yee [39]. Although there are several different

theories that can numerically describe electromagnetic propagation, the most complete is using Maxwell's equations with partial differential equations. Maxwell's equations are commonly understood to be the mathematical representation of the physical basis of lithographic imaging [9]. TEMPEST is a computer program that is used for solving Maxwell's equations with fully rigorous numerical solutions in a two dimensional domain. Utilizing massively parallel computers allows each simulating node or grid point to be assigned to a direct solution to Maxwell's equations. Computing in time-domain provides an accurate physical description of the solution at any instant. Convergence to a steady state solution is dependent on the physics of light scattering phenomena.

The fundamental concept behind electromagnetic waves is the existence of an electric field *E*, and magnetic field *H*, as a coupled vector. This simply indicates that light has both electric and magnetic fields that make up its properties as it propagates through a medium. An interaction between an electromagnetic wave on a material can give rise to an electric displacement *D*, magnetic induction *B*, electric current density *J*, and scalar quantity charge density ρ. Mathematically, these quantities can be explained by Maxwell's Equations. The differential form of Maxwell's Equations will be repeated here again:

$$\vec{\nabla} \times \vec{E} = -\frac{\partial \vec{B}}{\partial t} \tag{M.1}$$

$$\vec{\nabla} \times \vec{H} = \vec{J} + \frac{\partial \vec{D}}{\partial t} \tag{M.2}$$

$$\vec{\nabla} \circ \vec{D} = \rho \tag{M.3}$$

$$\vec{\nabla} \circ \vec{B} = 0 \tag{M.4}$$

Faraday's law of induction (Equation M.1) describes how an incremental change of a magnetic field can induce an electric field. Ampere's law (Equation M.2) describes how magnetic fields can be generated by electrical current and by changing electric field. Gauss' law for electricity (Equation M.3) states that the change in electric flux density is directly related to the charge density. Gauss' law for magnetism (Equation M.4) states that a magnetic flux through a loop must be a well-defined quantity. Each of these equations requires calculating solutions from the previous set. Convergence to a steady state solution is how Maxwell's equations are solved.

3.4.6 ELECTRIC AND MAGNETIC FIELD WAVE EQUATIONS

One of the main goals of introducing Maxwell's equations was to arrive at the electric and magnetic wave equations. The wave equation is given as [37]:

$$\nabla^2 \vec{A} = \frac{1}{v^2} \frac{\partial^2 \vec{A}}{\partial t^2}. \tag{28}$$

Also, vector operation gives a useful identity that:

$$\vec{\nabla} \times \left(\vec{\nabla} \times \vec{A} \right) = \vec{\nabla} \left(\vec{\nabla} \circ \vec{A} \right) - \nabla^2 \vec{A}. \tag{29}$$

First analyzing the electric field wave equation, the differential form of Faraday's law will be used (Equation M.11). If the fields are assumed to be sufficiently smooth, taking the curl on both sides gives the following:

$$\vec{\nabla} \times \left(\vec{\nabla} \times \vec{E} \right) = \vec{\nabla} \times \left(-\frac{\partial \vec{B}}{\partial t} \right) = -\frac{\partial \left(\vec{\nabla} \times \vec{B} \right)}{\partial t}. \tag{30}$$

The curl of a magnetic field and divergence of an electric field have already been determined to be Equation M.2 and M.3, respectively. Therefore, combining Equations

29 and 30 and plugging in Gauss' law for electricity and Ampere-Maxwell law, the following is obtained:

$$-\frac{\partial\left(\vec{J}+\frac{\partial\vec{D}}{\partial t}\right)}{\partial t}=\vec{\nabla}(\rho)-\nabla^2\vec{E}$$
$$\frac{\partial\vec{J}}{\partial t}+\frac{\partial^2\vec{D}}{\partial t^2}=\nabla^2\vec{E}-\vec{\nabla}(\rho).$$

In a material of interest that has no current or charge, the above equation simplifies to the wave equation for electric fields:

$$\nabla^2\vec{E}=\frac{\partial^2\vec{D}}{\partial t^2}. \tag{31}$$

A similar exercise obtains the wave equation for magnetic fields:

$$\nabla^2\vec{B}=\frac{\partial^2\vec{H}}{\partial t^2}. \tag{32}$$

Equations 31 and 32 describe how an electric and magnetic field propagates as a wave through a medium, a relatively simple outcome after understanding Maxwell's equations.

3.4.7 YEE'S ALGORITHM

Yee's algorithm was introduced in 1966 as a means to solve Maxwell's equations with finite differential equations [39]. Computer simulations such as TEMPEST utilize this algorithm because of its clever selection of geometry for spatial sampling of electric and magnetic field vector components. The basis of the FDTD numerical analysis for Yee's algorithm is simultaneously solving for the magnetic and electric fields associated with Maxwell's equations. While Yee's algorithm was developed in 1966, the algorithm

continues to be of great use because of its simplicity yet robust method for numerical solutions.

In Cartesian coordinates, Equations M.1 and M.2 are equivalent to the following scalar equations:

$$-\frac{\partial B_x}{\partial t} = \frac{\partial E_z}{\partial y} - \frac{\partial E_y}{\partial z} \tag{33a}$$

$$-\frac{\partial B_y}{\partial t} = \frac{\partial E_x}{\partial z} - \frac{\partial E_z}{\partial x} \tag{33b}$$

$$\frac{\partial B_z}{\partial t} = \frac{\partial E_x}{\partial y} - \frac{\partial E_y}{\partial x} \tag{33c}$$

$$\frac{\partial D_x}{\partial t} = \frac{\partial H_z}{\partial y} - \frac{\partial H_y}{\partial z} - J_x \tag{33d}$$

$$\frac{\partial D_y}{\partial t} = \frac{\partial H_x}{\partial z} - \frac{\partial H_z}{\partial x} - J_y \tag{33e}$$

$$\frac{\partial D_z}{\partial t} = \frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y} - J_z \tag{33f}$$

The above six partial differential equations forms the basis of the FDTD numerical analysis for Yee's algorithm that accounts for the electromagnetic wave interactions in a three-dimensional space. Before looking into the general three-dimensional case, it is beneficial to simplify to a two and one dimensional case for better understanding the Yee algorithm.

3.4.8 MAXWELL'S EQUATIONS IN TWO-DIMENSIONS

Assume a modeling dimension that is infinite in the *z*-direction with no changes in shape. If the incident wave is uniform along the *z*-axis, then all partial derivatives with respect to *z* must be zero. Equations 33a-f simplify to the following:

$$-\frac{\partial B_x}{\partial t} = \frac{\partial E_z}{\partial y} \tag{34a}$$

$$\frac{\partial B_y}{\partial t} = \frac{\partial E_z}{\partial x} \tag{34b}$$

$$\frac{\partial B_z}{\partial t} = \frac{\partial E_x}{\partial y} - \frac{\partial E_y}{\partial x} \tag{34c}$$

$$\frac{\partial D_x}{\partial t} = \frac{\partial H_z}{\partial y} - J_x \tag{34d}$$

$$\frac{\partial D_y}{\partial t} = -\frac{\partial H_z}{\partial x} - J_y \tag{34e}$$

$$\frac{\partial D_z}{\partial t} = \frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y} - J_z \tag{34f}$$

The above equations can be designated into two sets of field components as transverse-magnetic mode and transverse-electric mode with respect to z (TM_z and TE_z, respectively). Since the TM_z mode involves only one component of the magnetic field in the direction of interest, Equations 34a, b and f are designated as the TM_z in two dimensions. With similar reasoning, Equations 34c, d and e are designated to be the TE_z in two dimensions. In the TE_z mode, the electric field orientation is in the plane that is perpendicular to the z-axis or in plane with the surface of the structure.

3.4.9 ONE-DIMENSION WAVE EQUATION AND MAXWELL'S EQUATIONS

Further simplification of no electromagnetic field excitation or geometry variation in the *y*-direction, Maxwell's equations simplify to the following:

$$\frac{\partial B_y}{\partial t} = \frac{\partial E_z}{\partial x} \tag{35a}$$

$$\frac{\partial B_z}{\partial t} = -\frac{\partial E_y}{\partial x} \tag{35b}$$

$$\frac{\partial D_y}{\partial t} = -\frac{\partial H_z}{\partial x} - J_y \tag{35c}$$

$$\frac{\partial D_z}{\partial t} = \frac{\partial H_y}{\partial x} - J_z \tag{35d}$$

Now taking the partial time derivative of Equation 35a gives

$$\frac{\partial^2 B_y}{\partial t^2} = \frac{\partial E_z}{\partial t \partial x},$$

which simply requires a substitution of the partial space derivative of Equation 35d as shown below:

$$\frac{\partial^2 D_z}{\partial t \partial x} = \frac{\partial^2 H_y}{\partial x^2}$$

Therefore referring to Equation 5 and 6, the linearity of the system allows to combine to give the following result:

$$\frac{\partial^2 H_y}{\partial t^2} = \frac{1}{\mu\varepsilon}\frac{\partial^2 H_y}{\partial x^2}$$

Since $c = 1/\sqrt{\mu\varepsilon}$, the final result for the magnetic field in the y direction is given as

$$\frac{\partial^2 H_y}{\partial t^2} = c^2 \frac{\partial^2 H_y}{\partial x^2}, \tag{36a}$$

and applying the same analysis provide the result for the electric field in the z-direction to be:

$$\frac{\partial^2 E_z}{\partial t^2} = c^2 \frac{\partial^2 E_z}{\partial x^2}. \tag{36b}$$

Equation 36b as shown above is a one-dimensional scalar wave equation for an electric field in the z-direction. Therefore, Equations 34 provide a means to mathematically explain propagating waves of electric and magnetic field components that travel at the speed of light.

3.4.10 ILLUSTRATION OF YEE'S ALGORITHM

Yee's algorithm was developed to help model the set of finite differential equations for Maxwell's time-dependent curl equation systems shown in Equation 28. This is in contrast to solving the electric and magnetic fields separately using a wave equation. The three dimensional illustration of Yee's algorithm with its *E* and *H* components is shown in the figure below. The electric and magnetic field components are arranged in the cubic unit cell of Yee's space lattice such that both *E* and *H* have a surrounding circulating component. This allows simultaneously calculation of the differential and integral form of Maxwell's equations.

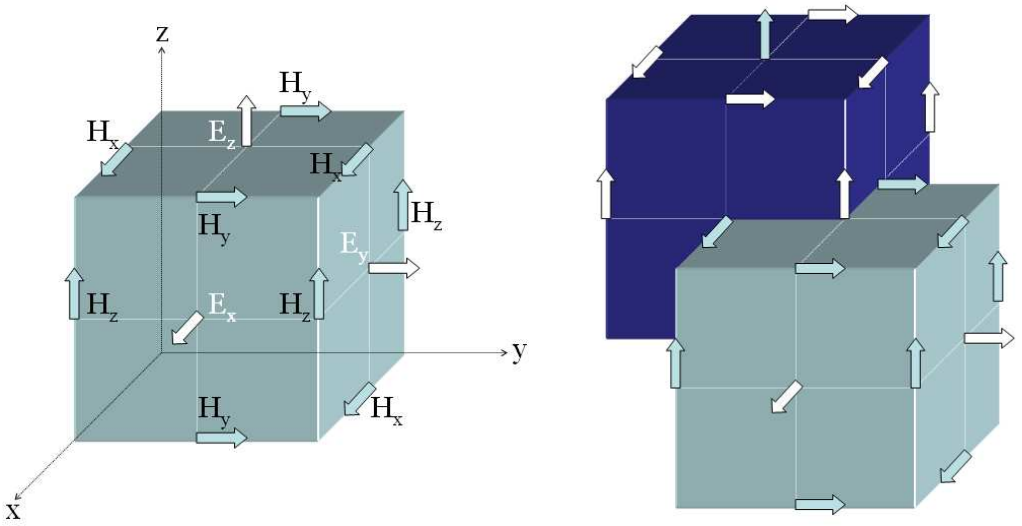


Figure 21. Three-dimensional illustration of Yee's algorithm showing *E* and *H* components in a cubic unit cell of Yee's space lattice [39].

Another benefit to Yee's algorithm is the time based structure of the leapfrog arrangement. Every *E* computation is completed and stored in memory for a particular time point using previously calculated *H* data. This cycle continues on with the *H* computation until the time-stepping is finished. Visually, this process can be understood by referring to the figure below.

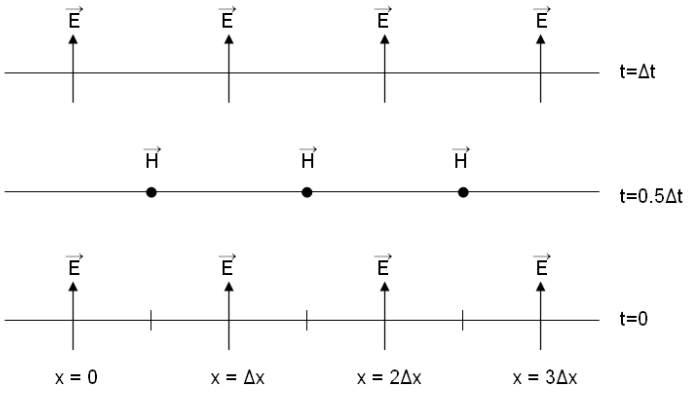


Figure 22. Leapfrog space-time step arrangement for Yee's algorithm in a one-dimensional wave propagation [39].

Panoramic Technology® combines the leapfrog space-time step arrangement, cubic unit cell, and Yee's algorithm to calculate Maxwell's equations. These methods are used for the rigorous TEMPEST simulations to accurately model advanced lithography applications.

3.5 THREE SUB-MODEL LITHOGRAPHIC IMAGING SYSTEM

Panoramic Technology® uses a three sub-model lithographic imaging system to simulate the electromagnetic wave nature of light within an imaging system. The three sub-models are the mask/object, imaging system, and resist model as shown in Figure 23. An example simulator flow will be used to help understand the simulation structure of Panoramic Technology®.

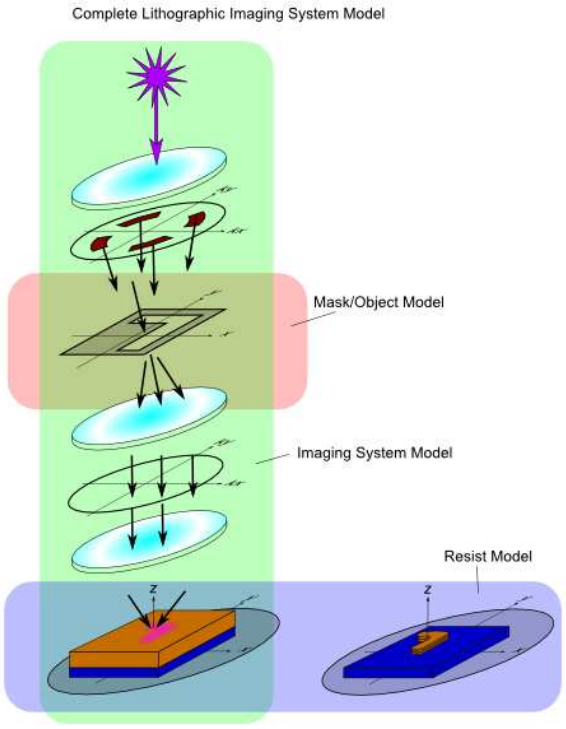


Figure 23. The three sub-model lithographic imaging system of Panoramic Technology®: mask/object, imaging system, and resist model [41].

In the simulation software, the three sub-model lithographic imaging system is represented by three separate simulator flows, each representing the respective mask/object, imaging system, and resist model, as shown in Figure 24. As an example, numbers from the preliminary simulation investigation discussed in chapter 6 will be used. A custom-defined mask is used to define the non-symmetric block mask with x_1 and x_2 at a given illumination wavelength. These user-defined parameters are used in a simulator that will calculate the scattered orders with a designated pitch. In this work, a Kirchhoff thin mask approximation is used as a mask simulator. The thin mask approximation treats the mask as a 2D object without any thickness [32]. The calculated scattered orders are then gathered into the imaging simulator along with input parameters such as NA and σ. Again, the simulator requires a calculation approximation such as Abbe's source integration method [42]. The aerial images from the defined pitch and input parameters are then calculated into the full physics resist simulator to output resist profiles with given resist parameter inputs.

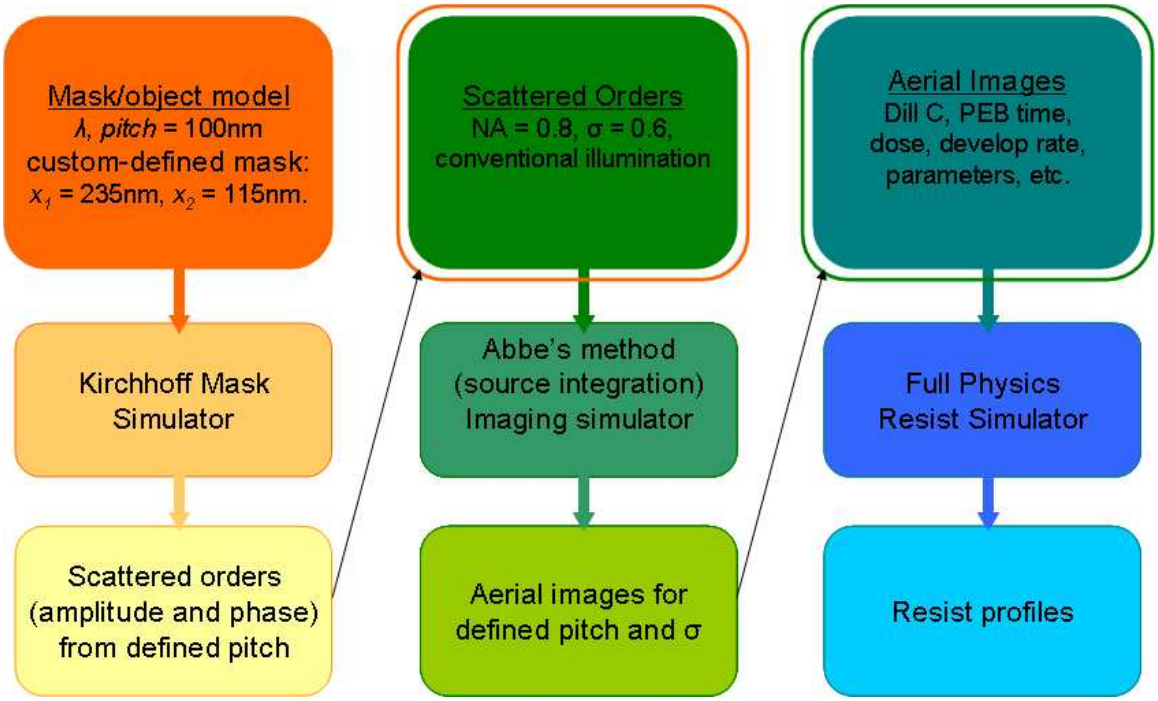


Figure 24. Example simulator flow of the three sub-model lithographic imaging system – (a) mask/object, (b) imaging system, and (c) resist model of Panoramic Technology® [41].

In general, Panoramic Technology® follows an input to simulator to output simulation structure. The generic simulator flow is shown in Figure 25. The simulation input defines the "setup," where several parameters are defined to include multiple cases or "runs" in a batch. These user-defined input parameters are then calculated with a simulator that runs each case or "runs" in the batch. The simulator can be thought of as the calculator that is based on the different approximation models (such as Kirchhoff thin mask, Abbe's method, and full physics resist simulator) for the different sub-model lithographic imaging system model. Once the calculations are complete, multiple outputs for each "run" in the batch are completed to produce the simulation output [41]. The distinction between being able to run the three sub-model calculations separately or automatically is the basic difference between EM-suite and Hyperlith, respectively. EM-

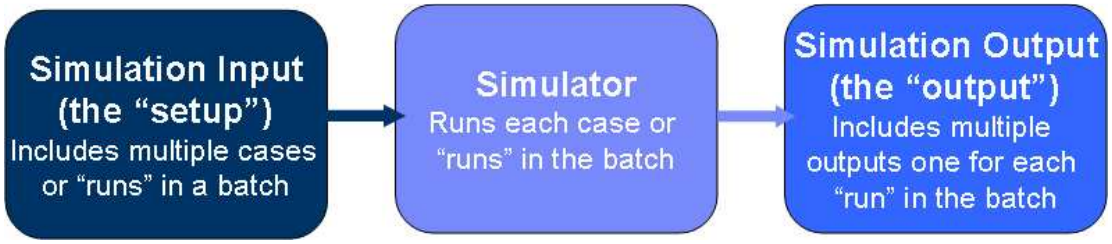


Figure 25. Generic simulator flow for the simulation structure of Panoramic Technology® [41].

suite allows a user to have better troubleshooting capabilities, while Hyperlith has more of a user-friendly approach.

The combination of the simulation structure of Panoramic Technology® and its FDTD calculations of Maxwell's equations defines an accurate modeling tool that describes EM wave propagation through its surroundings. One of the benefits to modeling work is its predictive capabilities. Another is the ability to implement low cost and quick methods of understanding and optimizing lithography processing. As lithography processing becomes increasingly difficult, the benefits for simulation work become clear.

3.6 Lithography Challenges of Scaling

The challenges of lithography scaling are meeting the CD and overlay tolerances while maintaining a reasonable cost of ownership. Solutions in addressing the challenges of scaling highly depend on the specific application.

CD and Overlay

Achieving CD control and overlay tolerances to design specifications are the two most difficult requirements in lithography processing [9]. As discussed in Section 3.3 because of tighter specifications, slight deviations from a target CD or overlay causes a

significant effect on fabricating transistors with higher yield. Deviations such as resist line edge roughness (LER) control to metrology and defect inspection in the single-nanometer regime can indirectly impact device degradation. Some lithographic processing deviations can be considered as defects because they are anomalies in the imaging process. The size of critical defects is generally accepted as half of the design rule structures [35]. Current challenges in lithography can range from bridging defects across one or more lines [36] to variations of resist patterns as a result of wafer topography. Innovative techniques are necessary to overcome these lithographic deviations.

There are several state-of-the-technology techniques at 248 nm for higher lithography performance. It is commonly understood that lithography at 248 nm illumination is a very mature technology [43]. Improvements in RETs such as exposure lens, NA enlargement and elaborate OPC and phase-shift masks designs resulted in better resolution [44]. OPC makes intentional design alterations to the mask that accounts for a blurring effect limited by the specific NA projection optics. Phase-shift masks are engineered in different material and thicknesses to utilize the phase shift of EM waves to induce better exposure intensities at the wafer level [32]. In general, intentional mask modifications enhance contrast. High contrast resists such as CAR and the addition of multi-layer resists, and anti-reflective coatings also enhance image quality. These techniques are used to extend the incumbent optical lithography technology.

Cost of Ownership

The alternative techniques are still in the development stages because the cost of ownership (COO) is still not optimized to the point where the industry can make a complete change [5]. Therefore, there are still efforts to extending the existing optical lithography techniques. Cost is the foundation to what drives the competitive nature of the semiconductor industry. Cost of tools, materials (such as resists, gas, other liquids, etc.), and masks are examples of economic factors that are used to determine the technology used for semiconductor processing. In general, the total cost of processing wafers is based on COO given as the following equation [22]:

$$COO = \frac{C_{fixed} + C_{\mathrm{var}iable} + C_{yieldloss}}{(TPT)(Y)(U)} \tag{27}$$

COO considers the fixed, variable and yield loss costs, equipment life, throughput rate TPT, ideal throughput yield Y, and utilization U. All of the cost variables in Equation 27 add to the total COO value. An example of fixed cost would be the cost of a new lithography equipment such as an EUV tool. Variable cost associated with just this new tool investment would include research cost going into efficient EUV optics, sources, masks, and photoresists. The cost of yield loss can be high for a newly developed tool, in comparison with a mature technology. The tool throughput, yield and utilization must ultimately be efficient enough to outweigh the substantial costs associated with this new tool [45]. The decision of implementing a NGL or an addition of a new processing is based on the COO and the risks involved.

Analyzing the theoretical yield aspects of COO and determining the benefits for a new investment can be done through simulation. Comparing the relative capabilities of

the new and old processes is dependent on resolution. Smaller CD result in more transistors per device that results in better chip technology available to the market. Simulation work is a valuable resource because of its ability to simulate essentially a limitless array of processing conditions without incurring processing cost. If a certain technique shows significant improvement that can prove the benefits of COO, then the simulation work can be a tool to verify the need for an addition of a process complexity. In this work, the process complexity of dBARC processing is analyzed to determine the benefit for improving the variations of resist block patterns as a result of wafer topography

CHAPTER 4

METHODS

Although the block mask dimensions are considered as non-critical, the impact of block mask CD variations to device output is far from an insignificant matter. The block mask CD variations have shown significant challenges to current device processing. The challenge investigated in this work was how specific lithography processing and its variation on block mask CD influenced the threshold voltage distribution.

4.1. BLOCK MASK LITHOGRAPHY

The block mask lithography and simulation was based on the illustration (not drawn to scale) as shown in Figure 26. The block mask geometries are defined by the *x* and *y*, which is also the total CD of the block mask in the *x* and *y* direction, respectively. Since the *x*-direction is perpendicular to the gate and the *y* is parallel to the gate, the terms length and width, respectively, will be used interchangeably. This dimension can be an opening or a block in the resist that defines the different halo implant treatments for NMOS and PMOS as discussed previously in Chapter 2. The green lines are the polysilicon conductors (PC) that define the gate structures. The term PC topography is introduced because not all of the PC lines are considered to be the gate. In fact, the fourth PC line from the left is the only gate structure in this test investigation. The

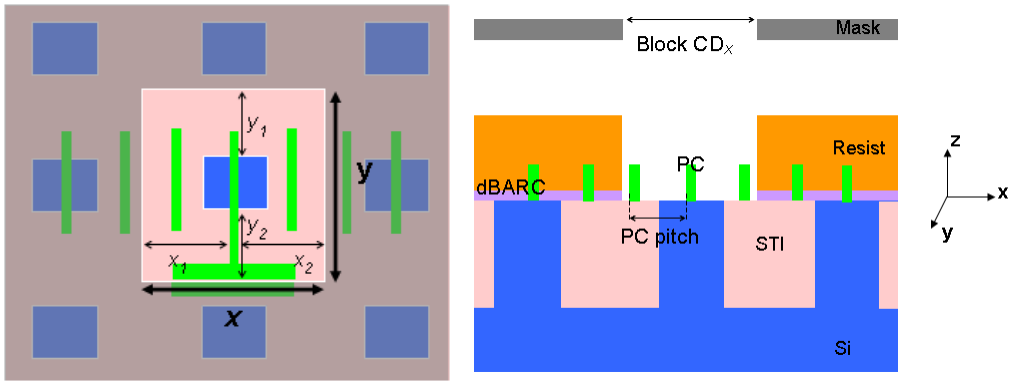


Figure 26. An illustration (not drawn to scale) of (a) an aerial view of the block mask lithography and (b) the cross-sectional view of the simulation domain. The PC topographies are shown as green lines, Si active regions as blue regions, STI as pink regions, resist as orange, dBARC as purple, and the block mask as gray, which are all input parameters.

existence of these PC structures can result in possible shadowing effects to the halo implant. The PC structures along with its pitch and dBARC, resist, STI, active regions, and block mask geometries are all input parameters in the simulation to accurately develop a physical model.

There is a clear distinction in the scanning electron micrographs (SEM) between different processing parameters of wavelength exposure and SLR versus DBARC layers. The micrograph below in Figure 27 shows a block mask opening, where the underlying PC topography are visible. The light vertical lines are the PC line topographies. The defined block mask seems to show more corner rounding when the wavelength used is 248 nm compared to 193 nm. A simplified visual representation is shown in Figure 30,

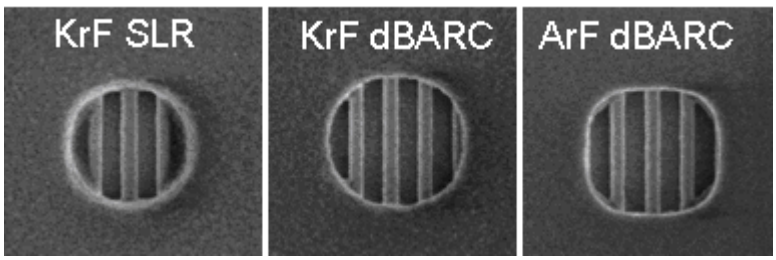


Figure 27. Micrograph showing an aerial image of a macro from a block mask opening with the different processing parameters [2].

where 248 nm wavelength shows the scenario in (c). The KrF SLR micrograph seems to show a dark blur around the edges of the block opening. This is the resist residual remaining around the corners between the wafer topography resulting from insufficient exposure intensity in the corners. In an ideal scenario, the SEM of the block mask geometry should show a rectangle, as represented by the scenario in Figure 30(a). The addition of dBARC processing complexity enables a lithography processing closer to the ideal situation. By visual inspection, the most beneficial processing choice for least variation of defined to exposed geometry is to utilize 193 nm wavelength with a dBARC layer.

The use of a dBARC layer sharpens the corners of the block mask and promotes a more ideal resist profile for halo implant. The resist block of KrF dBARC versus SLR in Figure 28 shows improvement in the edge integrity of a resist profile. The red measurement along the edges of the resist block shows the roughness of the edge. The reduction of the roughness measurement on the edge of the dBARC resist block shows that the edge integrity is improved. Without the dBARC processing, the rough corner edges, represented by Figure 30(b), decreases the actual resist block area at the top and increases the halo implant. The resist block SEM again shows the benefit to a dBARC processing.

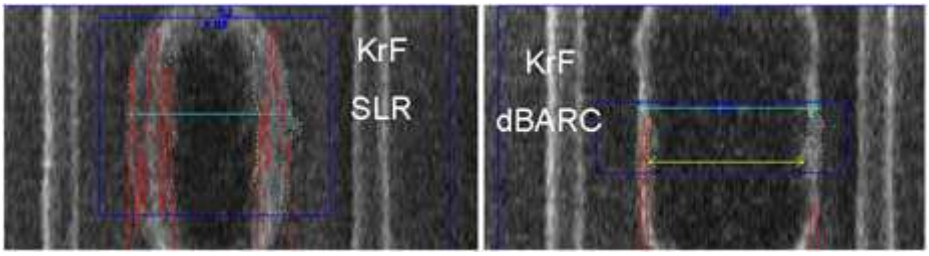


Figure 28. A micrograph showing the corners of the resist showing less rounding as the illumination wavelength is decreased from 248 to 193nm [2].

4.2. BLOCK MASK EXPERIMENTAL DESIGN

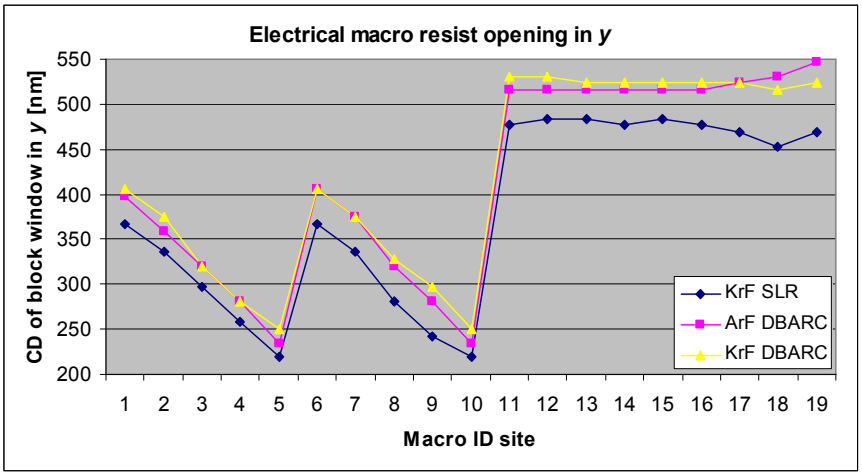
The experimental design used to calibrate the simulation model has intentional block mask dimensional changes to investigate the issue of block mask CD variation. The x and y dimensions are changed to x_1 and x_2 and y_1 and y_2, respectively, so that sensitive dimensions are not revealed. Although the active width and gate length values are not presented, the importance is the relative size different between the block mask geometry (~200 nm) to current wafer topography dimensions of about 20 nm [7]-[9], [43], [45]. In the current technology node, the ability to print a 200 nm geometry is considered non-critical.

Table 3. Different block mask defined geometries at macro ID sites with outlines for the different intentional geometry combinations [2].

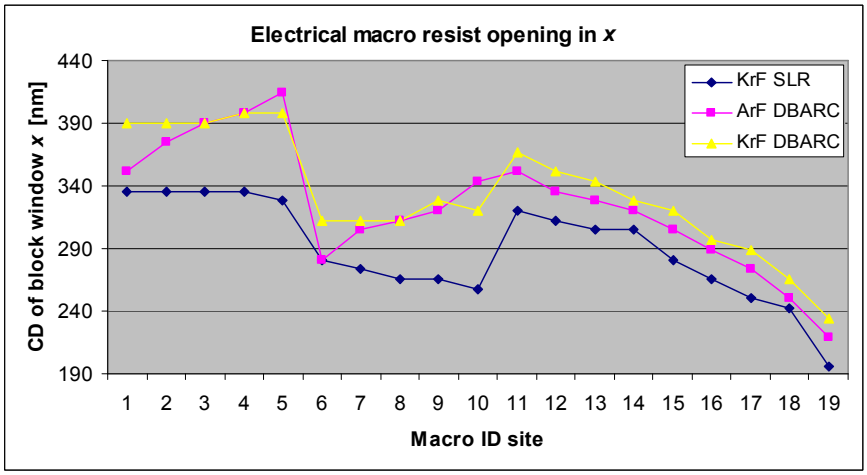
Simulation Defined Geometry				
Site	x1 [nm]	x2 [nm]	y1 [nm]	y2 [nm]
1	175	175	40	40
2	175	175	20	20
3	175	175	0	0
4	175	175	-20	-20
5	175	175	-40	-40
6	140	140	40	40
7	140	140	20	20
8	140	140	0	0
9	140	140	-20	-20
10	140	140	-40	-40
11	165	165	100	100
12	160	160	100	100
13	155	155	100	100
14	150	150	100	100
15	145	145	100	100
16	135	135	100	100
17	130	130	100	100
18	120	120	100	100
19	100	100	100	100

The block mask geometry variations are shown in Table 3 where each of the 19 treatment combinations corresponds to different macro ID test sites. The outlines of the green (1-5), purple (6-10), and orange borders (11-19) show a set of intentional variations. The first five have a constant x dimension and symmetric in x_1 and x_2, while the symmetric y dimensions are decreased. Sites 6-10 are similar to 1-5 except with a

smaller *x* dimension to represent a more aggressive dimension impinging on the width of the active region. The final nine geometries outlined in orange borders have a constant *y*-dimension with a decreasing symmetric *x*. These combinations are designed to investigate the influence of block mask CD variation with respect to the length of the active region. The measured data of the block mask geometries are shown in Figure 29, with three processing conditions of KrF (248 nm) SLR, KrF dBARC and ArF (193 nm) dBARC.



(a)



(b)

Figure 29. Measured data of block mask CD window in (a) *y*-dimension and (b) *x*-dimension with KrF SLR and ArF/KrF dBARC processing parameters [2].

These measured CD values were the basis to determining the validity of the simulation model used in the investigation. The root-mean-square error (RMSE) is a method to estimate the standard deviation of noise. Noise in this case can be interpreted as the difference between the model and the measured. A lower RMSE value on the order less than five nanometers is required to be confident on the model's accuracy and therefore its ability for predictive capabilities.

For modeling accuracy, simulation dose values need to match processing dose values. In Hyperlith, dose and exposure intensity is an arbitrary unit. The nominal dose was determined by a convergence by assigning an "anchor" or base macro dimension as the reference structure. The reference structure has large dimensions of one micron in the x and y so that there are no interference effects of the block mask and the wafer topography. The reference dose was then kept constant through the other block mask geometries.

The test site combinations of the different block mask geometries were electrically measured to investigate the effect on device output. In particular, it is the halo implant that is impacted by the block mask CD variation. In general, there are three

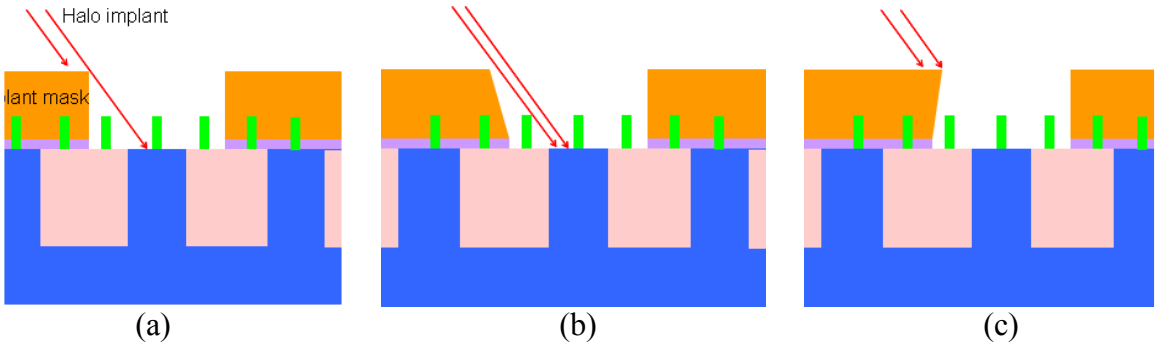


Figure 30. An conceptual illustration showing the different resist profiles that are related to an (a) ideal, (b) larger, and (c) smaller block mask CD and its associated halo implant effect.

scenarios that can be considered that can occur between a block mask CD variation and an incoming halo implant, as shown in Figure 30. The first scenario (a) is the ideal case when the resist profile is vertical and all the halo implant in the active region is as designed. The second scenario (b) is when the block mask CD is larger than designed and therefore increasing the amount of halo implant. The final scenario (c) shows a smaller block mask CD than designed that results in a lower amount of halo implant.

The block mask CD combinations were shown to have an influence on threshold voltage distribution. Variability chart for L_{nom} V_{Tlin} (Figure 2) will be reproduced here as Figure 31 to show the device output of different block mask geometries. Table 3 shows the different data sets relating to the intentional variations made to the block mask geometry. The one data point deviation between Table 3 and Figure 31 is from the "anchor" structure. In general, the first macro ID sites represent a relaxed *x* dimension with decreasing *y*-dimensions (green), followed by an aggressive *x*-dimension with decreasing *y*-dimensions (purple), and finally a constant *y*-dimension with decreasing *x*-dimensions (orange). Each data set has variations in one dimension and uniformity in the other to investigate block dimensions that impinge on their respective *x* and *y* dimensions

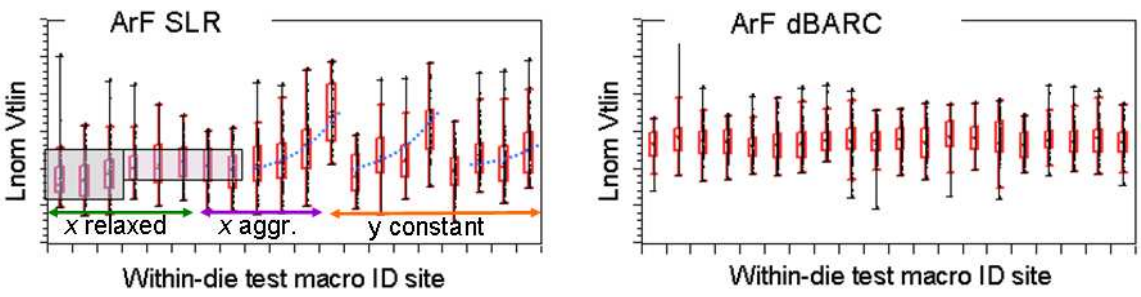


Figure 31. Graph of threshold voltage of ArF SLR (left) and ArF dBARC (right) processing with different within-die test macro ID sites. The macro ID sites for ArF SLR are separated into three data sets corresponding to the intentional variations shown in Table 3 [2].

of the active regions. The comparison between the ArF SLR and ArF dBARC threshold voltage clearly shows a difference in measured data. The macro ID design sites for the ArF SLR shows a systematic variation of threshold voltage, indicating a block lithography performance dependence on site location. The dBARC processing on the other hand improves the block mask lithography, resulting in a small random variation. Although the threshold voltage graph has several underlying mechanisms that may have influenced this outcome, the improvement shown with a dBARC over SLR processing gives strong interest to study how block mask CD variations effect threshold voltage.

4.3. GOAL OF THIS STUDY

The challenge of simulating a lithography system is to accurately predict and model the behavior of photoresists and other chemistries with various imaging and processing conditions. The accuracy can be improved by developing better physical and chemical models using calibration techniques. Calibration is determined by comparisons to input data such as CD measurements. The goal of this work was to develop an accurate physical model that simulates the measured characteristics of the block mask lithography, which has been shown to influence the device electrical performance.

CHAPTER 5

SIMULATION AND MODEL CALIBRATION

As stated in Chapter 4, the goal of this work was to develop an accurate physical model that simulates the measured characteristics of the block mask lithography. Panoramic Technology® is the simulation software that will be used to support the goal of this work. EM-Suite and Hyperlith simulations developed by Panoramic Technology® uses TEMPEST and FDTD algorithms to solve for Maxwell's equations. The versatile applications of EM-suite and Hyperlith range from defining complex 3D photomask and wafer topographies, constructing feature dimensions using variables and formulas, visualizing 2D cut-planes or 3D viewing capabilities, to calculating and plotting results to analyze multiple lithographic effects.

The first objective required to meet the goal was to use Panoramic Technology® simulation software to model CD variation of block mask lithography. The block mask geometries that are used for the calibration input for the simulation are from KrF SLR and ArF/KrF dBARC processing parameters. The specific dimensions values will be discussed in the next chapter. The measured data values were the target values used to analyze the modeling capability of representing observed behaviors.

The second objective that will support developing an accurate physical model was investigating the intensity profiles of the simulation domain as PC pitch is decreased.

Device scaling can result in undesirable effects of resist residuals in between PC lines because of limited light exposure related to the wavelength of the illumination source and the spatial frequency. The resist residuals remain at around 25% of the PC height.

5.1 Preliminary Simulations

Before the simulation and analysis work, preliminary simulations will be presented in this section to give a general relationship to how exposure intensity profiles are related to simulated geometries. The basic setup to the following simulations is having x_1 and x_2 values that are not symmetric as shown in Figure 32 with specific values of 235 and 115 nm, respectively. The x_1 and x_2 values can define a region for a block window, represented by the filled in shapes, or a block feature, shown as shaded shapes. The illumination wavelength is 193 nm. The vertical intensity profiles at x = 250, 350, 450, and 550 nm, or in between the PC lines will be analyzed. The lateral intensity profile will be located at 25% of the PC height.

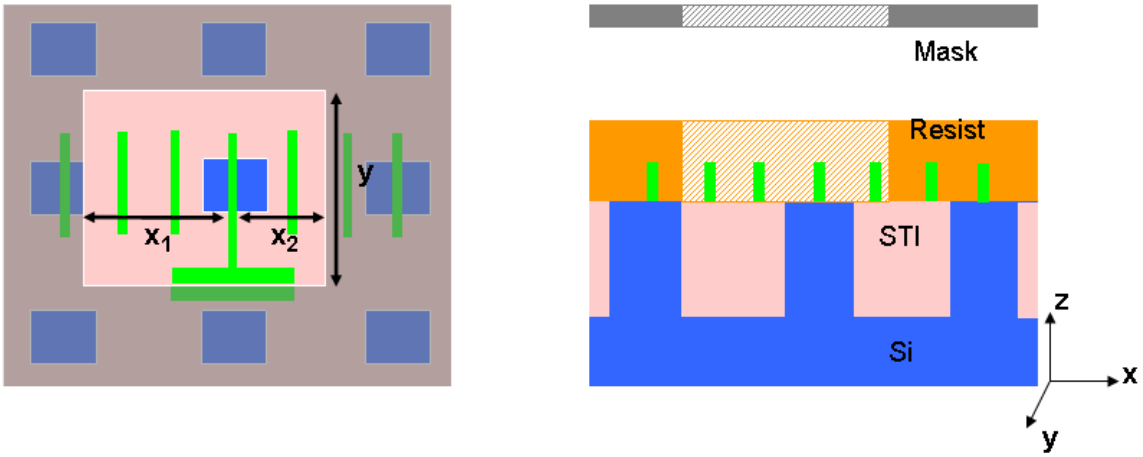
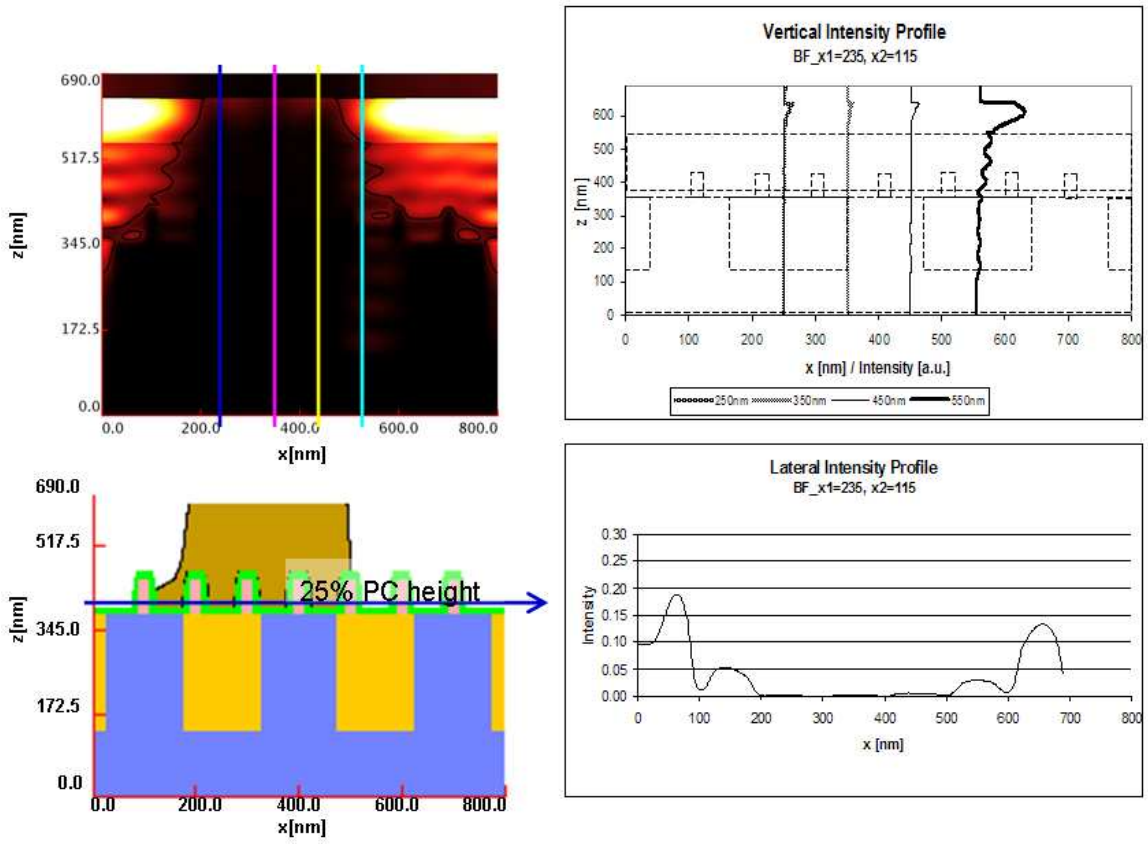


Figure 32. An illustration of the block mask geometry defined by *x1* and *x2*, shown on the left, for a possibility of printing a block window (colored), or a block feature (shaded) shown on the right.



(a) (b)

Figure 33. A preliminary simulation of a block feature mask: (a) simulation output of aerial image intensity profiles and the various vertical cutline locations of x = 250, 350, 450, and 550 nm (upper image) with its corresponding simulation domain defining the block feature mask x_1 and x_2 (lower image), and (b) the vertical intensity profile overlapped with the simulation domain (upper image) and the lateral intensity profile cutline at 25% of the PC height.

The result of the preliminary simulation of the block feature mask is shown in Figure 33. Standing waves are verified in the simulation output of aerial image intensity profiles and the vertical intensity profiles. The simulation domain input shown in the lower image of (a) shows that wafer topographies cause an undesirable effect to the resist profile. At the preliminary simulation pitch of 100 nm, the left corner of the resist block lips onto the first PC structure. This effectively prints a larger CD block feature structure.

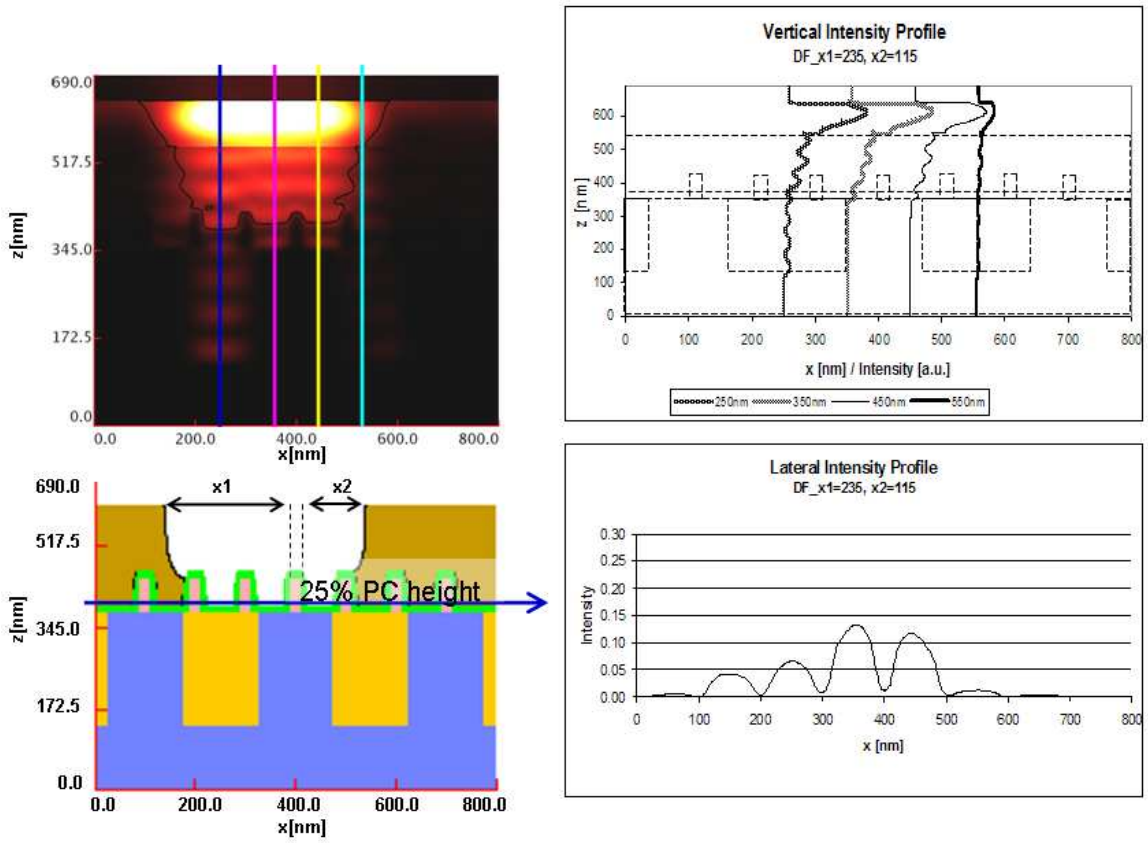


Figure 34. A preliminary simulation of a block mask window: (a) simulation output of aerial image intensity profiles and the various vertical cutline locations of x = 250, 350, 450, and 550nm (upper image) with its corresponding simulation domain defining the block mask window x_1 and x_2 (lower image), and (b) the vertical intensity profile overlapped with the simulation domain (upper image) and the lateral intensity profile cutline at 25% of the PC height.

The preliminary simulation for the block mask window shows similar results of standing waves and CD variations, shown in Figure 34. Instead of a straight resist profile that would show the same x_1 and x_2 dimension at the resist interface, the resist edge curves out toward the edge of the PC topography. This effectively causes the block mask CD to decrease. While these CD variations shown in the preliminary simulations occur near the top of the resist, resist residue can occur at any edges that light exposure can be limited. The intention of analyzing the intensity profile at 25% PC height is to investigate the issue of resist residue at the corner of a PC structure.

5.2 Block Mask Modeling

The purpose of block mask modeling is to accurately characterize block mask performance related to CD variation. The CD variations in current processing show that differences can cause problematic issues to device behavior. All modeling work requires calculations that can be simplified or be rigorous in nature. The complexity of the calculations in the simulation work presented in this work is related to the addition of process enhancement techniques. In particular, the work includes the wafer topography, device related geometries, and OPC mask. To investigate the benefits of simulation complexity, the modeling work will be split up into three separate models shown in Figure 35.

The first model is the complex scenario with both an OPC mask and wafer topography. The second model is simulated without any OPC mask but with wafer

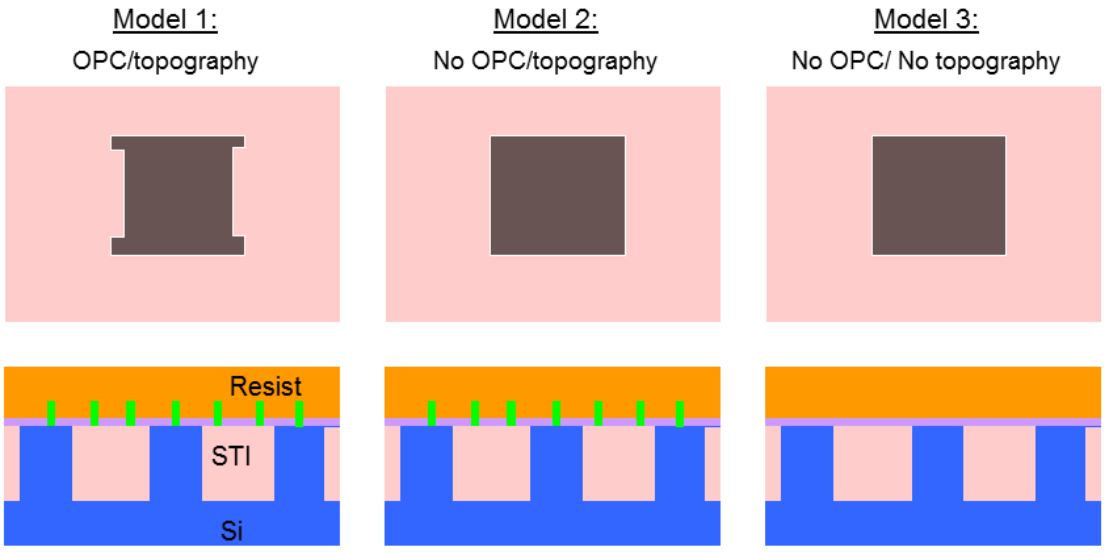


Figure 35. The three block mask modeling: 1. OPC with wafer topography geometries, 2. no OPC with wafer topography geometries, and 3. no OPC with any wafer topography geometries.

topographies of PC lines. The third model is a simulation without any OPC or wafer topography geometries. All three models will be compared to each other with each of the different processing parameters and separated into *x* and *y* block CD measurements.

The three simulated models are compared with the measured block mask CD as a percent difference with the measurement as the target. The graphical analysis of the *y* CD block measurement, the dimension along the PC lines, is shown in Figure 36. For example, negative values indicate a smaller geometrical output in the simulation compared to the measured CD. The first model with OPC with topography is shown as a solid triangle with bolded solid-lined error bars. The second model with no OPC with topography is shown as an open circle with dashed error bars. The third model with no OPC and no topography is shown as an "x" with solid-lined error bars. The target value of zero represents the measured CD block mask geometries shown in Figure 29. The scale between KrF dBARC and ArF dBARC is kept consistent to show the improvement

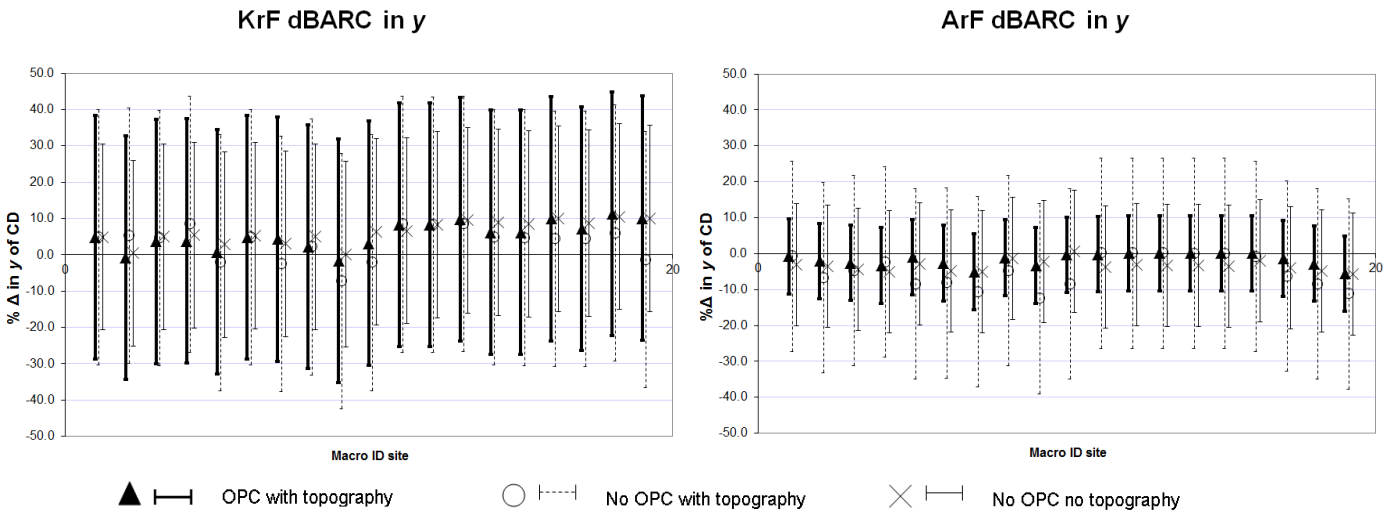


Figure 36. Graphical analysis of the percent difference between the *y* block mask CD measurements to the simulated values, with a slight offset in the macro ID sites. The first model of OPC with topography is represented as a solid triangle with dark solid-lined error bars, the second model of no OPC with topography is represented by a circle with dashed error bars, and the third model of no OPC with no topography is represented by an "x" with light solid-lined error bars.

of error variation using the shorter wavelength with ArF illumination. The variation in KrF dBARC does not show any significant difference between the different models. On the other hand, using the third model in ArF dBARC simulations shows an improvement with a significant decrease in the data variation.

All graphical details are kept consistent with the previous setup to make similar analyses for the *x* block mask CD variations that are perpendicular to the PC lines, shown in Figure 37. The error is slightly better with the KrF than the ArF. Both cases show an improvement in using the complex model of both OPC and topography compared to that of the no OPC with topography case. The benefits between the third model and the first are not as significant.

There are distinct differences between the graphs shown in Figure 36 and Figure 37. The variation between simulated and measured CD was expected to be larger for the *x*-dimension than the *y*-dimension because of the topographical effects that are changing

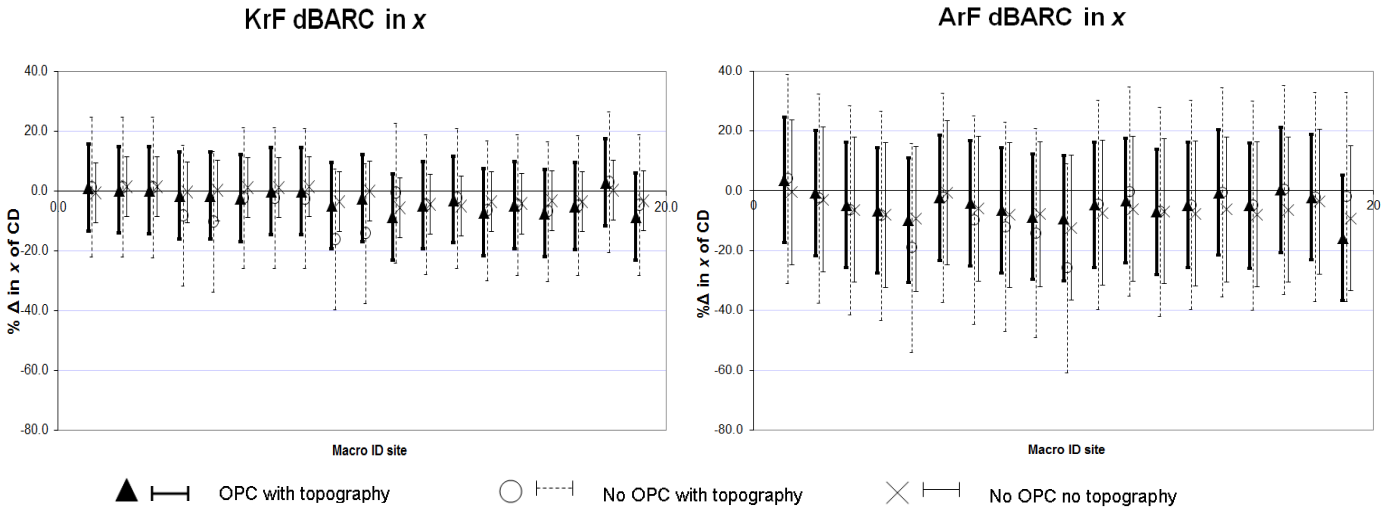


Figure 37. Graphical analysis of the percent difference between the *x* block mask CD measurements to the simulated values, with a slight offset in the macro ID sites. The first model of OPC with topography is represented as a solid triangle with dark solid-lined error bars, the second model of no OPC with topography is represented by a circle with dashed error bars, and the third model of no OPC with no topography is represented by an "x" with light solid-lined error bars.

perpendicular to the PC lines. The larger scale range in Figure 36 seems to indicate otherwise. However, the variation improvement shown using ArF illumination seems to follow the expected trend that shorter wavelength is related to better resolution. This variation improvement in Figure 37 is not as significant. In fact, it seems to show a slight advantage with KrF processing. The inclusion of the third model is to have a base simulation that does not have any modeling complexities. In most cases, the simulation for the third model seems to match the other two models. It may even show that certain cases, the third model shows the best comparison to the CD block mask. This neglects the fact that the third model is an unrealistic scenario and therefore cannot accurately represent a measurement. The important outcome to take out of this is that the fact that the base simulation shows reasonable correlation with measured but further simulation complexities are necessary for an accurate modeling of the block mask CD variations.

Since the block mask CD variation is a routine measure of model accuracy, the RMSE quantifies the validity of the modeling behavior as shown in Figure 38. A couple of observations can be made concerning the simulation modeling. First, when comparing samples with no OPC models, (shown as the first and second group in Figure 38), topography adds a significant amount to the RMSE, except the simulation parameter of KrF in *y*. This implies a significant induced variation to the overall process with an inclusion of wafer topography. Second, when comparing samples with topography (Group 1 and Group 3 in Figure 38), OPC significantly improves the RMSE. Generally, all processing parameters except for KrF in *y*, shows an improvement of at least 10 nm in CD variation (RMSE) with the addition of modeling complexity. In fact, the improvement of CD variation in ArF in both *x* and *y* dimensions is about 15 nm. This

shows a relatively well-behaved model that will be used to characterize a current lithography issue.

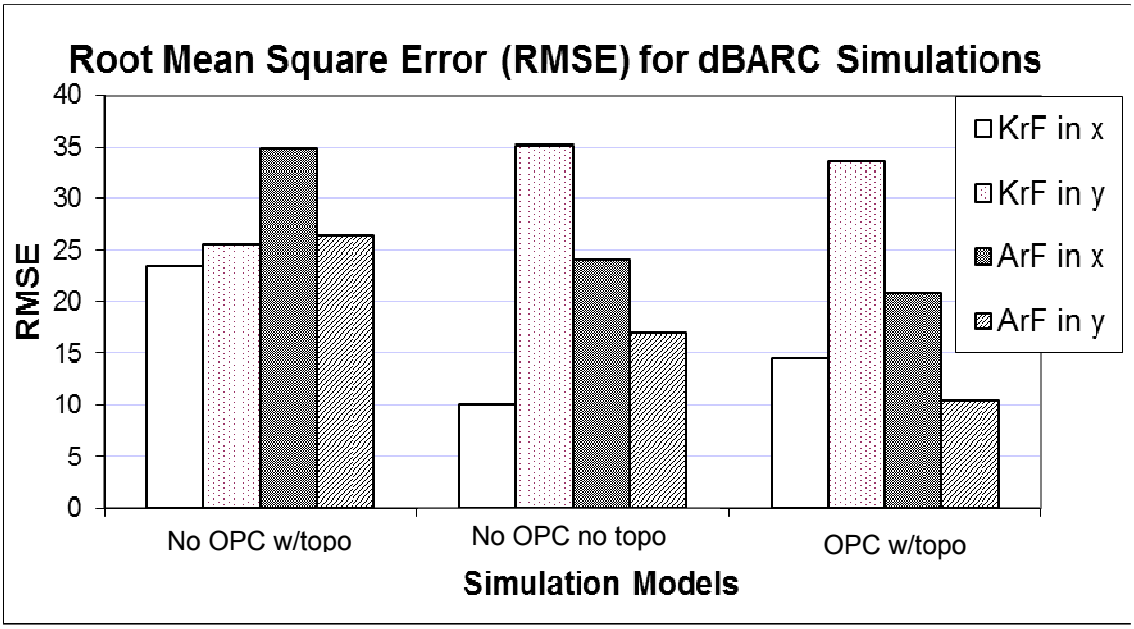


Figure 38. RMSE bar graph of the different models for the KrF and ArF dBARC parameters in *x* and *y* block mask simulated CD.

5.3 PC PITCH SCALING EFFECT ON INTENSITY

An OPC with topography model will be implemented to simulate the ArF exposure intensity profile as the PC pitch is decreased and making a comparison between dBARC and SLR processing. The basic setup is using 193 nm wavelength and previous defined geometries relative to a PC pitch variation from 1000 nm (an isolation case) to 40 nm (a dense pattern case). The block mask window will remain constant. Exposure intensity will be analyzed at two locations along the wafer at a 25% PC height. Simulated exposure intensity values will be taken at 5 nm feature proximity (open triangle) and the middle of the PCs (solid square), as shown in Figure 39. Each point is represented by its respective shapes that will be consistent with following graphs.

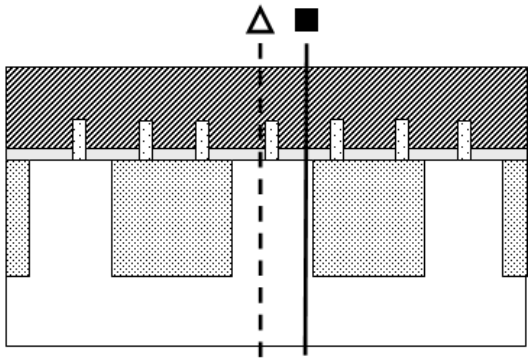


Figure 39. An illustration of two location of the simulated exposure intensity at 25% of the PC height. The 5 nm proximity to the PC feature is represented by an open triangle, while the central location between the PCS is represented by a solid square.

The simulated output of the exposure intensity at the center and 25% of PC height for a 193 nm SLR is shown in Figure 40. The independent variable increasing is spatial frequency or the inverse pitch. The pitch values simulated are 1000, 500, 250, 175, 150, 125 and 100:10:40 nm. The three regions of interest are represented by the three different slopes. The first three large pitch values show a relatively flat slope, as expected from underlying topography having no effect on exposure intensity. In the next region of points that indicates a positive slope, a regression analysis produces a slope of

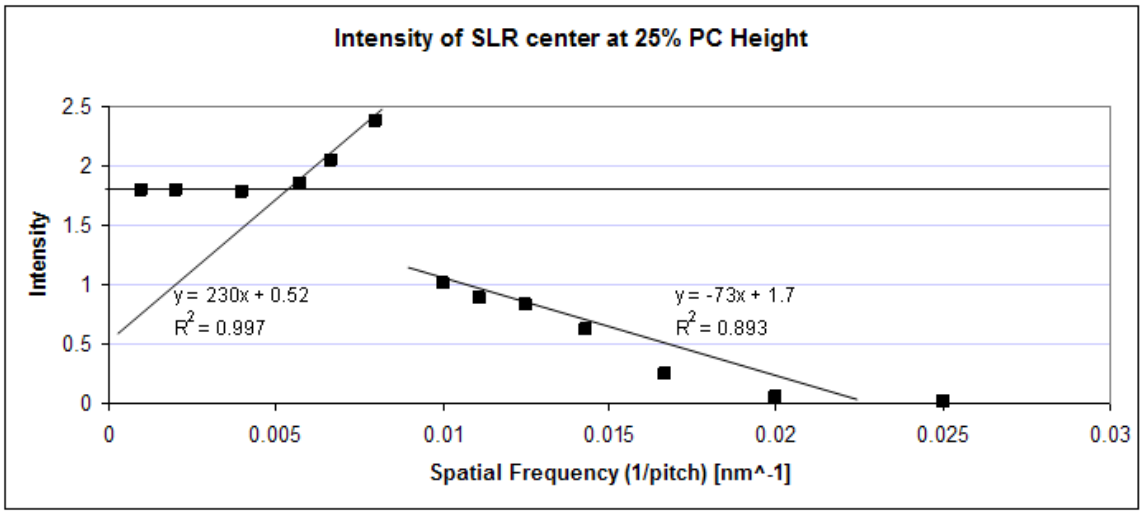


Figure 40. A graph showing the intensity variation as spatial frequency of a 193 nm SLR simulation located at the center and 25% of PC height.

230 with an R^2 value of 0.997. The following region shows an abrupt change of slope -73 with an R^2 value 0.893. The lower goodness of fit between the higher spatial frequencies indicates the necessity of a finer increment to see a better relationship between spatial frequency and exposure intensity.

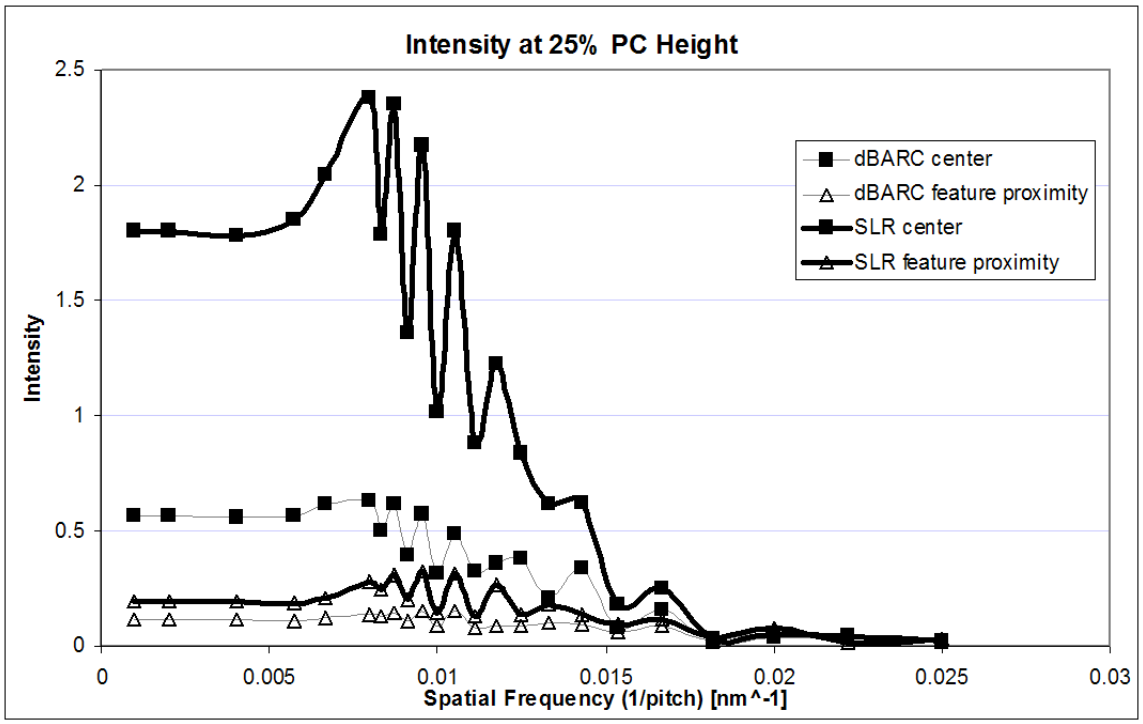


Figure 41. A graphical analysis of dBARC (gray line) and SLR (bold black line) simulation parameters output of exposure intensity with varying spatial frequency at the two locations of center of PC lines and 5nm feature proximity. A smoothed fit has been applied which reinforces the modulation effect, however it must be noted that the individual data points are not necessarily at the true local maximum or minimum intensity values.

With increased increments, the graph in Figure 41 shows very interesting results of exposure intensity as spatial frequency is increased of the different processing cases of dBARC layer and a SLR. The dark solid lines represent the SLR simulation and the gray lines represent the dBARC simulation. Although these lines are a smoothed fit, the modulation trend is valid. The highest and lowest points may not necessarily be represented in the graph, but a tighter spatial frequency increment shows that some

degree of modulation is present. Smaller increments will verify the highest and lowest points but in order to examine the spatial frequency effect on exposure intensity at 25% of PC height, this will be sufficient.

In all cases, the three regions shown in Figure 40 are consistent in Figure 41. The region of a relatively flat slope occurs in the large pitch values, while the positive slope is between 175 and 125 nm. The modulation region occurs in the 125:5:50 nm region. There seems to be a fourth region that shows another flat region of near zero intensity beyond 50 nm. Scaling at that magnitude will result in under-exposure and resist residue remaining in the corners.

The benefits to using a dBARC processing can be analyzed by taking the ratio between the center and feature proximity exposure intensity and comparing it with the SLR simulation. Figure 42 analyzes the ratio of center to 5 nm feature proximity exposure intensity for SLR and dBARC simulations. A higher ratio is related to higher

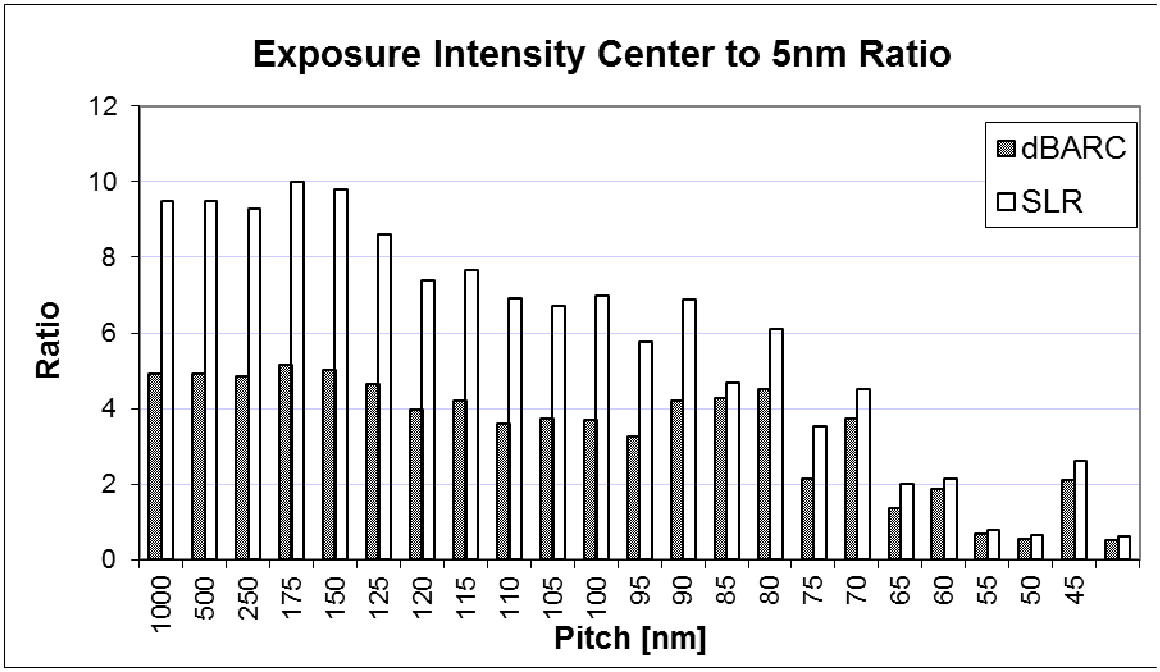


Figure 42. A bar graph comparing SLR and dBARC ratios between the center to 5 nm feature proximity exposure intensity with varying pitch values.

exposure intensity at the center compared to the PC edge proximity. This can be interpreted as higher light loss at the edges resulting in resist residues. All pitch lengths for SLR simulations in the bar graph show a larger ratio value. The OPC and topography modeling work shows greater light loss for SLR than dBARC processing.

CHAPTER 6

CONCLUSION

Finite Difference Time Domain simulation is used to relate block mask lithography issues of CD variation and light intensities to their undesired effect on device characterization. This work investigates the potential sources of and solutions for halo lithography variation. The block mask CDs are the geometries that define the implant region, therefore block mask CD variations can affect the designed implant dose. SLR processing shows a rounding effect to the resist defined on the wafer (Figure 27). If these geometries defined at the mask level are not patterned in the resist within design rules, the following processes can be negatively affected. Since a halo implant is at an angle, any resist residuals or corner rounding can cause implant species to impact the threshold voltage. Predictive simulation is valuable in understanding the fundamental origin of performance variation. Inaccuracies in block mask CD tolerances have shown to have significant effects on the electrical behavior of devices (Figure 2 and Figure 31).

Initial simulation work shows promising results on the ability to model processing results. The application of OPC and inclusion of underlying wafer topographies have shown an RMSE improvement of 10 nm for KrF in *x* and 15 nm for ArF in both *x* and *y* CD dimensions. For a first iteration modeling work, it shows reasonable modeling behavior. The benchmark value of RMSE to confidently use the model for predictive

purposes is less than 5 nm. This requires fine-tuning certain parameters in the model and recognizing its sensitivity. Additional calibration measurements from dBARC conformity will give a more accurate model as well. Consideration of other proximity effects such as active region reflection can also help develop a better model. In general, exposure intensity profiles along various locations (ideally infinite locations but realistically designed locations) give a better understanding on how varying wafer topographies influence patterned geometries. The results obtained from the fine-tuned model will ultimately help understand the fundamental causes of block mask CD variation. The target RMSE value not met in this modeling work determines the work still in its beginning stages. Although the mechanism behind block mask CD variation has not been fully understood, the first-iterated model shows promising results.

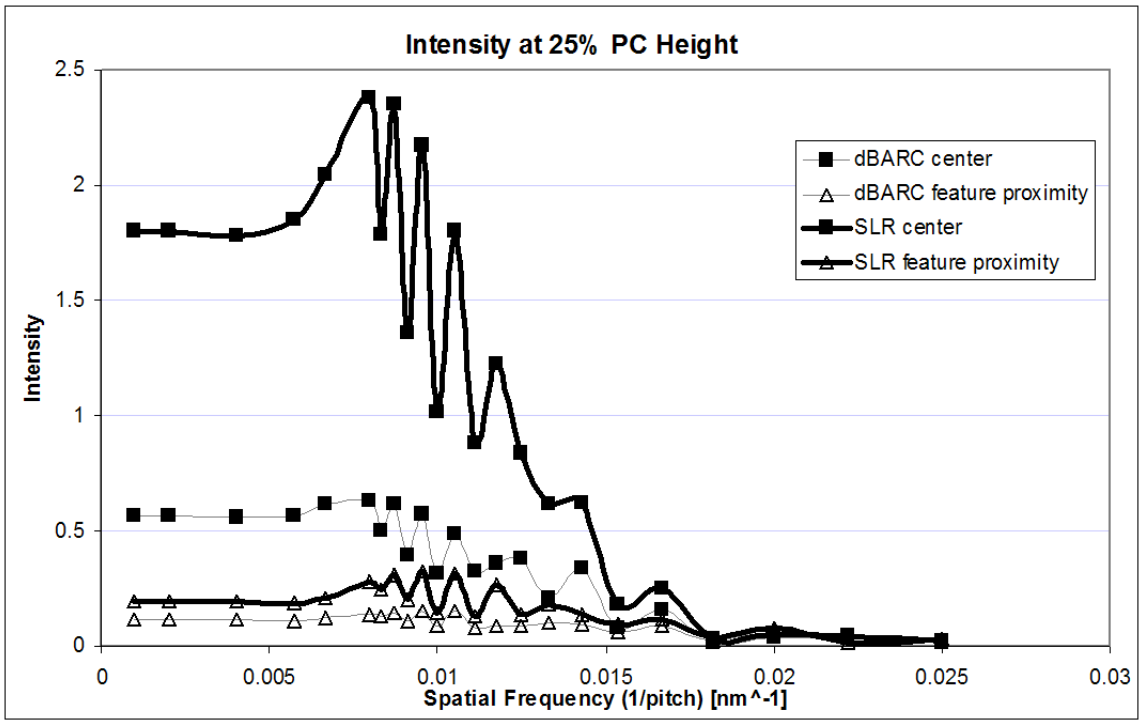


Figure 43. Graph of an exposure intensity variation with increasing spatial frequency showing the challenges the lithography industry faces for near-term Moore's law expectations (replication of Figure 41).

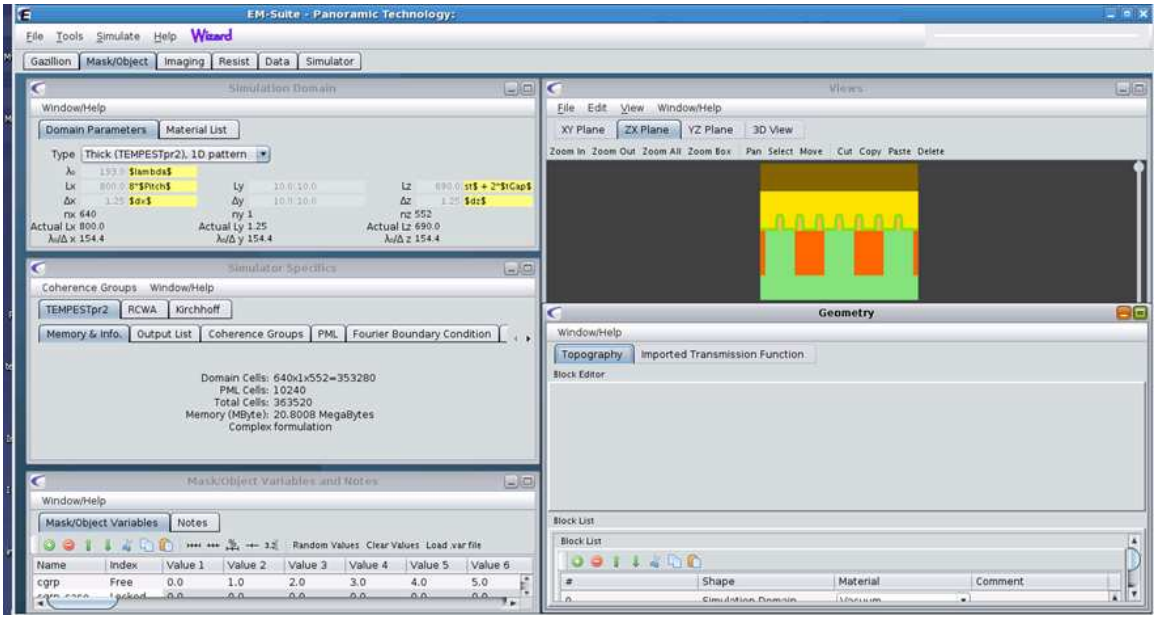
Even with a 15 nm RMSE value for the ArF OPC and wafer topography simulation model, the simulation is able to produce a meaningful result for exposure intensity variance with pitch frequency. Figure 41 is reproduced here to emphasize the significance demonstrated by this work. Even though block mask lithography is considered to be non-critical due to its large dimensions, and other lithography levels as critical, Figure 43 clearly shows the importance in underlying topography to the block mask lithography result. The modeling work shows that starting at a PC pitch of 125 nm, the exposure intensity exhibits swing curve modulation. The lithography step related to gate design cannot occur in this problematic range because of the necessity to have control in process latitude. A slight variation of the design pitch can lead to under or over-exposure, which ultimately leads to a failure in SCE control if solutions cannot be implemented effectively.

The development of an OPC with topography simulation have demonstrated the benefits to an ArF dBARC processing for improved exposure intensity and therefore related to minimizing reflection, reducing block mask CD variation, improving implant control, and improved device yield. However, as mentioned in Section 3.6, the ultimate decision is based on economics. Engineers require the skill set of simultaneously optimizing processing parameters related to device performance and COO. COO associated with current mask engineering is so high that it could limit technology scaling. The start-up cost for a mask is more than $10 million. Volume production cost of a mask-set for a 65 nm technology is close to $3 million while the cost for a 32 nm technology is about double [46]. Any process complexity is essentially an addition to cost. The implementation of an improved method, such as the dBARC processing, may

seem beneficial, but is at the same time, a new process. A new process requires investment including fixed, variable and yield loss. Benefits to throughput rate and yield may increase but the end goal is to maximize profit when marginal revenue equals marginal cost. The task may seem daunting, but in a semiconductor industry where there are so many unanswered questions, the range of research, projects, and accomplishments are limitless. Even a modest improvement in a non-critical block mask lithography process may ultimately make a significant contribution towards technological advancements.

APPENDIX A

The following image shows an example screenshot from EM-Suite, a software simulation package from Panoramic Technology®. TEMPEST is the simulator based on finite-difference time-domain (FDTD) calculations used in EM-Suite. Within the TEMPEST simulator, the input parameters that were defined in this work were the illumination source and type, projection optics parameters, custom mask layout design with the mask technology (thickness), calibrated dose, simulation domain dimensions, wafer geometries and layers, material parameters, full-physics resist model with wafer topography, and other necessary input parameters.



REFERENCES

[1] S.B. Felch, *et al.*, "90 nm device validation of the use of a single wafer, high-current implanter for high tilt halo implants," *Nuclear Instruments and Methods in Physics Research B*, vol. 237, pp. 53-57, 2005.

[2] S.J. Holmes, "Implementation of Reflectivity Control on Block Mask Lithography," *Proceedings of SPIE – Optical Microlithography XXV,* in press.

[3] G.E. Moore, "Cramming more components onto integrated circuits," *Electronics*, vol. 38, pp. 114-117, 1965.

[4] J. Lyman, "Optical Lithography Refuses to Die," *Electronics*, vol. 58, pp. 36-38, 1985.

[5] S. Wurm, "Lithography Development and Research Challenges for the $\leq$ 22nm Half-pitch," *Proceedings of the SPIE - The International Society for Optical Engineering*, vol. 7470 (11 pp.), 2007.

[6] I. Nishiyam, "Current status and perspective of EUV lithography," *Transactions of the Institute of Electrical Engineers of Japan, Part C,* vol.126-C, pp. 676-82, 2006.

[7] ITRS, "2010 Update: International Technology Roadmap for Semiconductors Lithography." Available: http://www.itrs.net/reports.html.

[8] ITRS, "International Technology Roadmap for Semiconductors Lithography 2009 Edition Executive Summary."

[9] ITRS, "International Technology Roadmap for Semiconductors Lithography 2009 Edition Lithography."

[10] A. Sekiguchi, Y. Matsumoto, and K. Okubo, "Study of Leaching Analysis of Immersion Resist Components using the WEXA-2 System," *Journal of Photopolymer Science and Technology*, vol. 23, pp. 427-33, 2010.

[11] P. Scholz, *et al.*, "A versatile design of solid immersion lenses in bulk silicon using focus ion beam techniques," *Conference Proceedings from the 35th International Symposium for Testing and Failure Analysis, ISTFA 2009,* pp. 119-25, 2009.

[12] H.Yaegashi, *et al.*, "Implementation of Double Patterning process toward 22-nm node," *Proceedings of the SPIE - The International Society for Optical Engineering*, vol. 7520 (9 pp.), 2009.

[13] J. Cameron, *et al.*, "Developable BARC (DBARC) Technology as a Solution to Today's Implant Lithography Challenges," *Proceedings of the SPIE - The International Society for Optical Engineering*, vol. 7972 (10 pp.), 2011.

[14] P. Innocenzi, "Patterning block copolymer thin films by deep X-ray lithography," *Soft Matter*, vol. 6, pp. 3172-3176, 2010.

[15] H.C. Pfeiffer, "Direct write electron beam lithography: a historical overview," *Proceedings of the SPIE - The International Society for Optical Engineering*, vol. 7823 (6 pp.), 2010.

[16] E. Slot, *et al.*, "MAPPER: High throughput maskless lithography," *Proceedings of the SPIE - The International Society for Optical Engineering*, vol. 7271 (8 pp.), 2009.

[17] T.G. Oyama, *et al.*, "Extendibility of EUV resists in the exposure wavelength from 13.5 down to 3.1 nm for next-generation lithography," *Proceedings of the SPIE – The International Society for Optical Engineering*, vol. 7972, pp. 797210 (6 pp.), 2011.

[18] M. Nastasi and J.W. Mayer, *Ion Implantation and Synthesis of Materials*, New York: Springer, 2006.

[19] R.S. Muller and T.I. Kamins, *Device Electronics for Integrated Circuits*, 2nd ed., New York: Wiley: 1986.

[20] J.D. Plummer, M.D. Deal, and P.B. Griffin, *Silicon VLSI Technology*, Upper Saddle River, NJ: Prentice Hall, 2000.

[21] S. Wolf, *Microchip Manufacturing*, Sunset Beach, CA: Lattice Press, 2004.

[22] S. Venkatesh and D.T. Phillips, "The SEMATECH cost of ownership model: An analysis and critique," Texas SCOE Texas A&M University.

[23] D. Meshulach, *et al.*, "Advanced Lithography – Wafer Defect Scattering Analysis at DUV," *Proceedings of SPIE – The International Society for Optical Engineering,* vol. 7638 (11 pp.), 2010.

[24] K. Ehara, "CD and Defect Improvement Challenges for Immersion Processes," *Proceedings of SPIE – The International Society for Optical Engineering,* vol. 7273 (12 pp.), 2009.
[25] S. Wolf, *Microchip Manufacturing*, Sunset Beach, CA: Lattice Press, 2004.
[26] W.J. Smith, *Modern Optical Engineering*, 3rd ed., New York: McGraw-Hill, 2000.
[27] L.S. Pedrotti, "Basic Physical Optics," in *Fundamentals of Photonics: Basic Physical Optics*, C. Roychoudhuri, Ed. Waco, Texas: SPIE Press Book, pp. 118-168, 2008.
[28] B.J. Lin, *Optical Lithography: Here is Why*, Bellingham, WA: SPIE Press, 2010.
[29] Reiser et al., "Novolak–diazonaphthoquinone resists: The central role of phenolic strings," *American Vacuum Society,* vol. 18, pp. 1288-1293, 2000.
[30] S. Franssila, *Introduction to Microfabrication*, Hoboken, NJ: John Wiley & Sons Ltd., 2004.
[31] K. Suzuki and B.W. Smith, *Microlithography Science and Technology*, 2nd ed., Boca Raton, FL: CRC Press, 2007.
[32] C. Mack, *Fundamental Principles of Optical Lithography*, England: John Wiley & Sons Ltd., 2007.
[33] K. Suzuki and B.W. Smith, *Microlithography Science and Technology*, 2nd ed., Boca Raton, FL: CRC Press, 2007.
[34] M.D. Levenson, "Wavefront engineering for photolithography," *Physics Today*, vol. 46, pp. 28-36.
[35] D. Meshulach, *et al.*, "Advanced Lithography – Wafer Defect Scattering Analysis at DUV," *Proceedings of SPIE – The International Society for Optical Engineering,* vol. 7638 (11 pp.), 2010.
[36] K. Ehara, "CD and Defect Improvement Challenges for Immersion Processes," *Proceedings of SPIE – The International Society for Optical Engineering,* vol. 7273, (12 pp.), 2009.
[37] D. Fleisch, *A Student's Guide to Maxwell's Equations*, New York: Cambridge University Press, 2008.
[38] R. Guerrieri, K.H. Tadros, J. Gamelin, and A.R. Neureuther, "Massively Parallel Algorithms for Scattering in Optical Lithography," *IEEE Transactions On Computer-Aided Design of Integrated Circuits and Systems*, vol. 10, pp. 1091-1100, 1991.
[39] K.S. Yee, "Numerical Solution of Initial Boundary Value Problems Involving Maxwell's Equations in Isotropic Media," *IEEE Transactions on Antennas and Propagation*, vol. 14, pp. 302-307, 1966.
[40] A. Taflove, and S.C. Hagness, *Computational Electrodynamics: The Finite-Difference Time-Domain Method*, 3rd ed., Norwood, MA: Artech House, inc., 2005.
[41] T. Pistor, *Panoramic Technology*, Berkley, CA, 1999.
Available: www.panoramictech.com
[42] D. Peng, P. Hu, V. Tolani, and T. Dam, "Toward a Consistent and Accurate Approach to Modeling Projection Optics," *Proceedings of SPIE – The International Society for Optical Engineering*, vol. 7640, (9 pp.), 2010.
[43] L. Knoll, *et al.,* "20nm gate length Schottky MOSFETs with ultra thin NiSi/epitaxial NiS_2 source/drain," *2011 12th International Conference on Ultimate Integration on Silicon (ULIS 2011)*, (4 pp.), 2011 .
[44] T.J. You, "Sub 100nm DRAM cell patterning results and relation with lens aberration at 248nm lithography era," *Proceedings of SPIE – The International Society for Optical Engineering*, vol. 5040, pp. 1327-34, 2003.
[45] S. Wurm and K. Kemp, "EUV Lithography," *Academie des Sciences. Comptes Rendus, Physique*, vol. 7, pp. 875-86, 2006.
[46] T. Jhaveri, *et al.*, "OPC Simplification and Mask Cost Reduction using Regular Design Fabrics," *Proceedings of SPIE – The International Society for Optical Engineering*, vol. 7274, (8 pp.), 2009.